Developments in Measuring and Modeling Financial Volatility

Janus, P.

2012

document version
Publisher's PDF, also known as Version of record

Developments in Measuring and Modeling Financial Volatility

ISBN 978 90 361 0274 2

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **523** of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

VRIJE UNIVERSITEIT

Developments in Measuring and Modeling Financial Volatility

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M.Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Economische Wetenschappen en Bedrijfskunde
op donderdag 16 februari 2012 om 11.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Paweł Janus

geboren te Leszno, Polen

promotor: prof.dr. S.J. Koopman
copromotor: dr. C.S. Bos

To my family

Acknowledgements

Completing the PhD degree is a challenging task. When I started in September 2008, I thought I have plenty of time. But I (very) soon realized that the research process from scratch requires deep understanding of the area of your interest, which makes you question the knowledge and motivates to search for improvements. This is fun but also hard work. No doubt, this demanding process calls forth a supervision. I was fortunate to have excellent guidance over the whole project. I am indebted to my supervisors Siem Jan Koopman and Charles S. Bos without whom this thesis would have never been completed. It was a great pleasure to conduct research under their supervision.

I thank Siem Jan Koopman. His expertise in time series analysis was invaluable for my research. He learned me how to critically review the knowledge we have and how to climb on the learning curve. He provided me with many critical remarks on how to structurize and how to present the research output. Also his optimism and support were of great importance to me over the whole period. I thank Charles S. Bos. His experience in computational econometrics helped my data-intensive research be so much efficient. In my daily work, he often advised me on finding efficient computational solutions relevant for my research. Also, his broad experience in financial econometrics and his attention to detail had stimulating impact on my research skills. Thank you Siem Jan and Charles for your time and for being accessible over the last three years.

I would like also to thank André Lucas. We worked together on one project. He shared with me with many constructive comments regarding parametric modeling of volatility. I would like also to thank Peter Reinhard Hansen. I had opportunity to visit him at the Stanford University where we worked together on covariance modeling. His knowledge of all 'ugly facts' about high-frequency data and his expertise in financial econometrics provided me with many valuable insights.

A special thanks goes to all current and past members of the time series group here at the VU University Amsterdam and at the University of Amsterdam with whom I had a pleasure to work. Thank you for all insightful discussions over lunches, coffees and econometric workshop.

Back at the Tinbergen Institute over 2006-2008 I had pleasure to meet many great people from all over the world. I will not point out any individual names here as I fear I

would forget some important ones. I remember long nights fighting together to understand some complex models, econometric assignments or game theory (I am now familiar with the 'no pain, no game' concept). Looking back at these years I see that at the heart of the voyage lies the experience of the whole crew sailing the ship together. Thank you TI folks.

I could not have completed this thesis without the support of my family and family-in-law who convinced me to keep going. Completion of this project is largely to their effort and I am eternally grateful for the unwavering encouragement which was of great importance, particularly during hard times. I am particularly thankful for the love of my father and my mother, and for everything they gave me to get where I am today.

Finally, I would like to thank two most important people in my life, my wife Katarzyna and my son Konstanty, the best boy any father could ask for. Their confidence in me was the backbone of the success I have achieved. I am so much thankful for their patience and love. Now it is time for me to pay off and I look forward to many great years ahead of us.

Amsterdam, August 2011

Contents

Chapter 1

Introduction

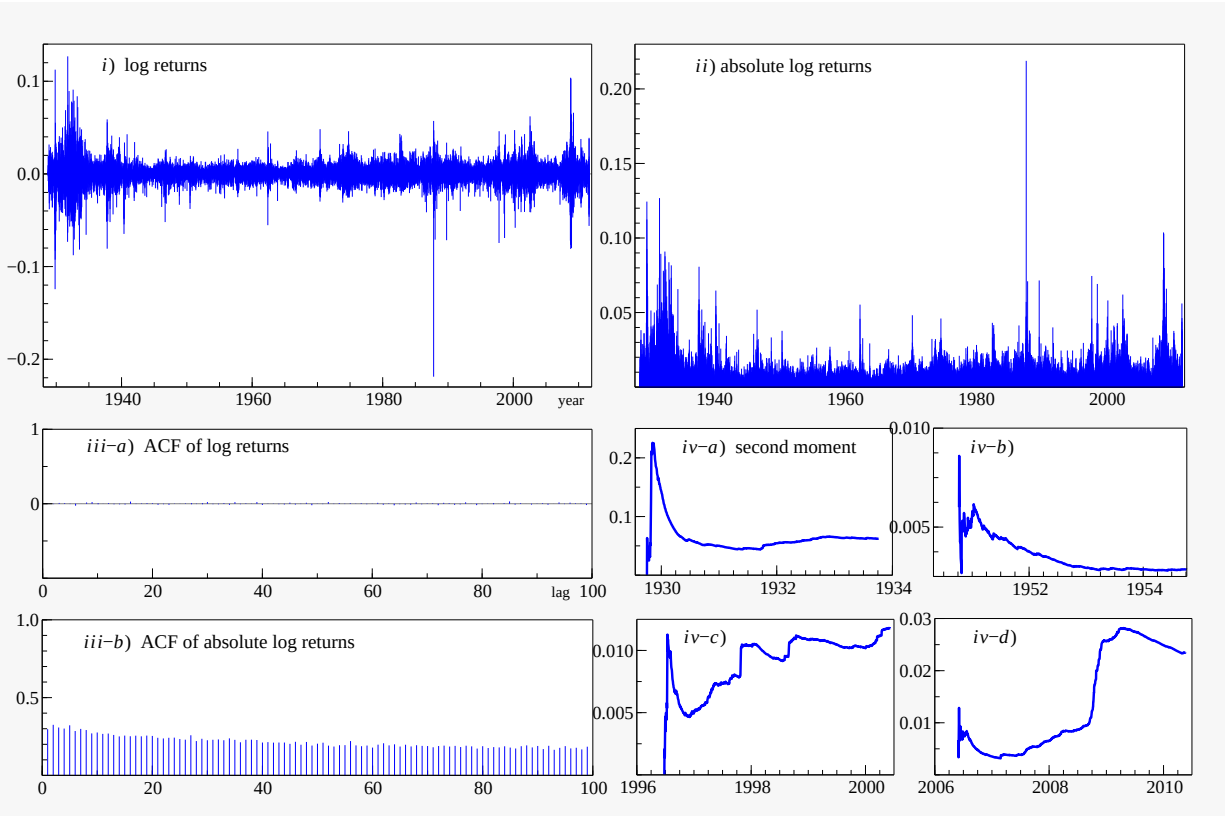
1.1 Motivation

In this thesis we present developments in measuring and modeling financial volatility. In financial econometrics, the term volatility is used to denote the measure of variation of asset price changes. Volatility describes the variation of a single asset, but when dealing with multiple assets then the multivariate volatility may also denote the whole dependence structure. The exact meaning of volatility as used here is specifically defined in each of the particular chapters of this thesis.

It is undisputed that financial time series have time-varying and persistent volatility, i.e. volatility exhibits different levels over time and clustering of episodes with low or high variation. For example, analyzing a time series of cotton price changes, Mandelbrot (1963, p.418) noted that "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes". One such example is illustrated in Figure 1.1, where we observe several periods when price changes of the Dow Jones Industrial index display high variation during certain intervals of time, and low variation at other times; see panels *i*) and *ii*). A statistical consequence of this is that, while financial returns themselves are uncorrelated, absolute (or squared) returns display strong, positive and slowly decaying autocorrelation; compare panels *iii-a*) and *iii-b*) of Figure 1.1. The time-varying property of volatility has initially been diagnosed by sample second moment estimated with increasing sample. Mandelbrot (1963) found that the second sample moment of price changes does not level-off to any limiting value as more observations are included for its estimation but it rather varies in an absolutely erratic fashion; this effect is illustrated here in panels *iv-a*) to *iv-b*) for different starting points. Similar features of price variation have been observed in virtually all assets, from equities to foreign exchange rates and home prices.

The phenomenon of time-varying asset price variation has been explored extensively in

Figure 1.1: *Properties of returns on DJI index*



Note: Dow Jones Industrial Average (DJI) index. The sample period runs from October 1, 1928 to August 1, 2011. ACF stands for the autocorrelation function. The second sample moment in panels *iv-a*) to *iv-d*) is computed with increasing sample sizes from 1 to 1000 observations; see note of Mandelbrot (1963, Figure 2, p.396).

the literature. On the one hand, a number of economic theories have been proposed in an effort to explain the origin of time-varying volatility. The theories often justify this origin through the actions of competing agents who make decisions based upon their asymmetric information sets which form their beliefs about the fair value of underlying assets. A diversity of strategies of agents, heterogeneous arrival rates of information or different investment horizons may all, to some extent, explain the source of volatility clustering. On the other hand, a variety of statistical tools and time series models have been developed and applied to characterize time-varying volatility of financial assets. The aim of this thesis is to advance various aspects of measuring and modeling financial volatility.

The concept of volatility is of equal interest to both academic and business circles. All tradeable assets and instruments are exposed to financial market volatility. The finance industry is thus heavily dependent on reliable estimates of volatility, which is relevant from risk management to option or bond pricing and hedging. The volatility measures provide also insight into the risk-return trade-off and risk-adjusted returns which play a central role in portfolio allocation and performance evaluation. To conclude, volatility of

financial returns is fundamental to many issues in financial economics, and the importance of its accurate measurement and modeling can hardly be overstated. This importance was indeed highlighted by the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2003, which was partially awarded to Robert F. Engle "for methods of analyzing economic time series with time-varying volatility (ARCH)".

The next subsections provide a brief summary of standard models and methods used to analyze volatility of financial assets, to give a description of the landscape in which this thesis will elaborate some aspects of volatility estimation. Almost by definition, such a brief summary will not pay sufficient tribute to many authors in the field, as it will only provide a partial description of some methods; for a wider overview, see e.g. Andersen, Davis, Kreiß, and Mikosch (2009).

1.2 A General Setup for Financial Data Modeling

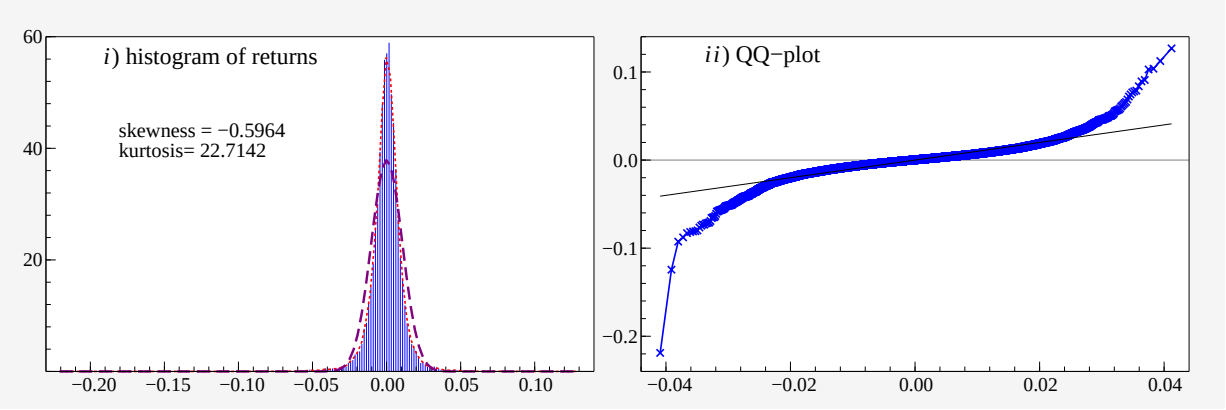
Financial data modeling starts with assumptions on a process for the evolution of asset prices. A common starting point is to let the underlying efficient logarithmic price process be characterized by a continuous-time diffusion process

$$\mathrm{d}X_t = \mu \mathrm{d}t + \sigma \mathrm{d}W_t, \qquad t \geq 0, \tag{1.1}$$

where μ denotes the drift term, $\sigma > 0$ denotes volatility, and W_t is standard Brownian motion. The standard diffusion process (1.1) assumes that successive log price changes, $X_{t+s} - X_t$ say, are independently and normally distributed random variables with variance proportional to the time difference s between prices. Because of tractability, the implied normality of asset price changes proves convenient in many aspects in financial economics, for example, Black and Scholes (1973) apply the normality assumption to derive their option pricing formula. Despite the fundamental importance of the stochastic process (1.1), it has become evident that it is not able to reproduce certain statistical features of financial returns. Already in the 1960s it was observed, notably by Mandelbrot (1963) and Fama (1965) among others, that empirical distributions of asset price changes are too *peaked* and *heavy-tailed* relative to samples from the normal distribution implied by (1.1). This fact is clearly illustrated here in Figure 1.2 for a series of daily returns on the Dow Jones Industrial index. The non-normality of asset returns is also evident from excess kurtosis and skewness that the diffusion process (1.1) does not account for. The peakedness and heavy-tailness of financial returns are principal features for virtually all financial assets and regardless of the frequency at which price data is sampled, though with a closer match to a normal density at low, monthly say, frequency.

The restrictive nature of the constant volatility diffusion process (1.1) has fostered development of more advanced models that promise a closer match with stylized facts regarding

Figure 1.2: *Density and QQ plot of returns on DJI index*



financial returns. One generalization considers continuous-time stochastic volatility models; see Hull and White (1987) and Scott (1987) for work in the context of option pricing. The continuous-time stochastic volatility models read

$$\begin{aligned} \mathrm{d}X_t &= \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t, \qquad t \geq 0, && (1.2) \\ \mathrm{d}\sigma_t &= f(\sigma_t, \mathrm{d}B_t), && (1.3) \end{aligned}$$

where, in general, the drift term μ_t and volatility σ_t are possibly time-varying. The dynamics of volatility σ_t has a functional form f which typically is autoregressive to capture persistence and is driven by stochastic innovation B_t, say a Brownian motion possibly correlated with W_t. The stochastic volatility diffusions are capable to reproduce fat tails typically observed in the empirical distribution functions of financial returns. The amount of excess kurtosis is controlled by the coefficients determining the functional form for stochastic volatility process (1.3), while the degree of skewness is controlled by the correlation between asset returns and stochastic volatility ($\mathbb{E}[\mathrm{d}W_t \mathrm{d}B_t]$), the so-called leverage effect. Stochastic volatility models however do not account for large discontinuous price movements, and may not be capable to resolve significant skewness often observed in the data; see Das and Sundaram (1999). For this reason, another generalization concerns the introduction of a jump component, see Merton (1976) who proposes to distinguish *normal* and *abnormal* changes in the stock prices. The jump model of Merton (1976) takes the following form

$$\mathrm{d}X_t = \mu \mathrm{d}t + \sigma \mathrm{d}W_t + \kappa_t \mathrm{d}N_t, \qquad t \geq 0, \tag{1.4}$$

where μ, σ are defined as in (1.1), κ_t is a random variable with mean $\mu_\kappa(t)$ and variance $\sigma^2_\kappa(t)$ with N_t being a counting process that represents the number of jumps in the price path up to time t. The importance of the jump component is apparent from the, possibly quite rare, occurrences of large price changes. For example, one such a substantial price

change attributed to the financial market crash of October 19, 1987 can be clearly seen in panel *i*) of Figure 1.1 and also in panel *ii*) of Figure 1.2. The level of excess kurtosis in the jump model (1.4) is determined by jump frequency and relative magnitude, while the degree of skewness can easily be governed by allowing for asymmetric jumps in the asset price process.

A number of studies illustrate that both generalizations discussed above appear to be vital to capture all the statistical aspects of asset price changes, and both stochastic volatility as well as jumps ought to be explicitly included in the asset price processes; see Das and Sundaram (1999). Consequently, combining stochastic volatility with jumps has now become standard, and the resulting model for the price process is given by

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t + \kappa_t \mathrm{d}N_t, \qquad t \geq 0, \tag{1.5}$$

with μ_t, σ_t as defined in (1.2), and with κ_t as defined in (1.4). The degree of skewness and excess kurtosis in the distribution of asset returns is determined by the parametrization of both stochastic volatility and jump processes.

From the econometric point of view, the stochastic volatility jump diffusion process (1.5) creates new challenges. It is important to measure and model time-varying volatility σ_t robustly to jumps in empirical data. But it is also of great interest to test for the presence of jumps (e.g. their number and timing) and to differentiate the price variation due to jumps from that to the continuous sample path as both have very different implications in financial economics. These two concerns outline effectively the central focus-point in the following chapters of this thesis.

1.3 Methods

The approaches to quantify financial volatility from empirical datasets can be categorized into non-parametric, semi-parametric and parametric. The non-parametric approach focuses on direct measurement of ex-post price variation over certain fixed time intervals. Within the parametric approach one constructs and exploits various functional forms for the underlying volatility of financial asset(s). Finally, the recently proposed semi-parametric methods make use of time series of ex-post measurements of volatility constructed from high-frequency data, and they apply standard time series techniques to model and forecast financial volatility.

1.3.1 Non-parametric volatility measurement

As noted already by Merton (1980), a simple but reasonable estimate for the variance term σ^2 in (1.1) is the sum of the squares of the logarithmic returns provided that asset prices

can be described by a diffusion-type stochastic process of the form (1.1). In particular, let us assume that price observations from (1.1) are observed over a certain fixed time interval $[0, T]$ at times $0 = \tau_0 < \ldots < \tau_n = T$. The log returns $\Delta X_{\tau_i} = X_{\tau_i} - X_{\tau_{i-1}}$ of the price process (1.1) are independently and identically normally distributed with mean $\mu\Delta\tau$ and standard deviation $\sigma\sqrt{\Delta\tau}$, where $\Delta\tau = T/n$. Merton (1980) suggested to estimate mean and variance as given by

$$\widehat{\mu} = \frac{1}{n\Delta\tau}\sum_{i=1}^{n}\Delta X_{\tau_i} = \frac{X_{\tau_n} - X_{\tau_0}}{T} \qquad \text{and} \qquad \widehat{\sigma}^2 = \frac{1}{n\Delta\tau}\sum_{i=1}^{n}\left(\Delta X_{\tau_i}\right)^2. \tag{1.6}$$

Note that the variance estimator in (1.6) differs from the maximum likelihood estimator as it is not centered around the mean. It can be shown that

$$\mathbb{E}[\widehat{\mu}] = \mu \qquad \text{and} \qquad \mathbb{V}ar[\widehat{\mu}] = \sigma^2/T, \tag{1.7}$$

and

$$\mathbb{E}[\widehat{\sigma}^2] = \sigma^2 + \mu^2 T/n \qquad \text{and} \qquad \mathbb{V}ar[\widehat{\sigma}^2] = 2\sigma^4/n + 4\mu^2\sigma^2 T/n^2. \tag{1.8}$$

The result in (1.7) illustrates that accuracy of the mean estimator as measured by its variance depends on the length of the fixed time interval T, and not on the number of observations n available within time interval $[0, T]$. This means that precise estimation of the mean return μ requires long spans of data, while sampling frequency determined by n is not relevant. Considering the variance estimator, it follows from (1.8) that it has a bias which however vanishes as n becomes large, in fact as $n \to \infty$, even within fixed time interval $[0, T]$. The resulting variance estimator is consistent and its accuracy depends on the number of observations n. The result of Merton (1980) implies that unlike estimation of the mean term, precise measurement of the variance can be obtained over fixed time intervals provided that the number of asset price observations is sufficiently large. The relevance of this result has become apparent with accessibility to data sampled at intraday high-frequencies.

1.3.2 Parametric models for volatility

Let Y_t denote the asset price on day t and let μ_t denote the conditional mean which is often imposed to be constant, i.e $\mu_t = \mu$. The demeaned log return is defined as $y_t = \log(Y_t/Y_{t-1}) - \mu$. The discrete-time stochastic volatility model for asset returns can be written as

$$y_t = \sigma_t\varepsilon_t, \qquad t = 1, \ldots, N, \tag{1.9}$$

where ε_t is an independently and identically distributed stochastic innovation with a standardized density function (Normal, Student's t etc.), and where $\sigma_t > 0$ denotes volatility

on day t. To model the time-varying volatility and to ensure its persistence, one typically considers some form of the autoregressive process for its dynamics as given by

$$\sigma_{t+1}^2 = \delta + \beta\sigma_t^2 + \alpha\eta_t, \qquad \delta, \beta, \alpha > 0, \tag{1.10}$$

or

$$h_{t+1} = \delta + \beta h_t + \alpha\eta_t, \qquad \text{with} \quad h_t = \log\sigma_t^2, \tag{1.11}$$

where, in general, δ, β and α are unknown coefficients, and where η_t is the innovation that drives the one-step ahead prediction of the volatility. The dynamic formulation in (1.11) is expressed for log volatility rather than volatility to ensure the volatility remains positive at all times.

It is common practice to model the (log) volatility as time-varying functions in the form of discrete-time generalized autoregressive conditional heteroskedasticity (GARCH) models of Engle (1982) and Bollerslev (1986). Another class of models that enables a time-varying variance is the discrete-time stochastic volatility (SV) model of Taylor (1986) and Harvey, Ruiz, and Shephard (1994). The nature of the innovation term η_t determines the class of volatility model. For example, in case we let η_t be a function of the current observation, $\eta_t = y_t^2 - \sigma_t^2$ say, we obtain a model that corresponds to the GARCH model. In case η_t is a stochastic disturbance term, the model is part of the SV class.

1.3.3 Semi-parametric models for volatility

While the GARCH and SV models - in their many variants - have proved successful in the past two decades, the increasing availability of high-frequency intraday data provides additional information about variation of asset price changes. In general, the main drawback of traditional GARCH and SV models is that they exploit only daily price observations to infer about the current level of volatilities or covariances of assets. Given that the non-parametric approach exploits intraday price observations to provide volatility measures that are consistent, approximately unbiased and with limited measurement error, there has been interest to integrate information from high-frequency data into modeling financial volatility.

Computation of ex-post measures of volatility from high-frequency data permits the use of more standard time series techniques for modeling and forecasting financial volatility. To illustrate this possibility more in detail, let us assume that we observe prices of some asset at intraday high-frequency times $0 = \tau_0 < \ldots < \tau_n = T = 1$, where the interval $[0, 1]$ represents a trading day. Having high-frequency price observations, we can compute $\widehat{\sigma}_t^2$ using (1.6) for each day t, with $t = 1, \ldots, N$. The resulting time series of volatility measures $\{\widehat{\sigma}_t^2\}_{t=1}^N$ can be modelled by an autoregressive formulation as

$$\log\widehat{\sigma}_t^2 = \delta + \beta_1 \log\widehat{\sigma}_{t-1}^2 + \ldots + \beta_p \log\widehat{\sigma}_{t-p}^2 + \epsilon_t, \tag{1.12}$$

where $\delta, \beta_1, \ldots, \beta_p$ are unknown coefficients, and where ϵ_t is the innovation term. The logarithmic transformation in (1.12) is again taken to ensure that volatility is positive at all times. Andersen, Bollerslev, Diebold, and Labys (2003) were the first to explore the use of the autoregressive model to analyze the time series of realized volatilities and they noted gains relative to fully parametric models.

1.4 Overview of the Thesis

The remainder of this thesis consists of four independent research papers and it quite naturally separates into two parts. Chapters 2 and 3 focus on direct volatility measurements using high-frequency data. Chapters 4 and 5 address modeling of financial volatility, in particular they aim to model the time-varying dependence of multiple assets.

In chapter 2 we discuss a non-parametric method to estimate the volatility at any time within a trading day using high-frequency price data. In particular, we consider the stochastic volatility diffusion process with jumps (1.5) and we aim to estimate volatility at any intraday time. The method is not hindered by the occurrence of jumps, microstructure noise, daily recurring patterns in the variance, or leverage effects. The resulting spot volatility path allows to track variation of prices at short intraday horizons, or to test for jump effects related to e.g. news announcements. We discuss a testing procedure designed to test for jumps in each of the price increments.

Chapter 3 focuses on robust estimation of daily price variation obtained from intraday data. Specifically, we propose a variance estimator based on quantiles of intraday price changes. In a combination with the standard non-robust moment-based measure of variance (1.6), the two estimators provide a sound platform for the construction of a test statistic designed to detect abnormal price changes over fixed time intervals. A new class of tests is developed and tested.

In chapter 4 we present a framework designed for analyzing the dependence structure of bivariate data. Particular attention is paid to capturing long-range autocovariances of volatilities and correlations. We illustrate the importance of the parametric assumptions about the observed data when constructing functional forms for both volatility and correlation in a similar way as outlined in (1.10)-(1.11). We show that the key to obtain reliable estimates of time-varying volatility and correlation is a robust formulation which mitigates the impact of jumps and of other outliers. This is a fully parametric approach.

Chapter 5 presents tools designed to model the covariance matrix of multiple assets. We illustrate how to incorporate intraday information for the modeling of covariance matrices of financial assets. The modeling framework is classified as semi-parametric because its idea corresponds to (1.12). The modeling approach in this chapter makes use of both low and high-frequency data in the form of measures containing signals about the current level

of volatilities and covariances.

Each chapter is self-contained, has its own notation and can be read independently. Each chapter is preceded with an abstract that highlights major developments and findings. We discuss why standard statistical methods may not be sufficient for providing good and reliable volatility estimates, and we point out where the methods of this thesis promise to provide better results. The novelty and contribution of the method are discussed separately in the individual chapters, where we relate our research results to existing alternative approaches.

Chapter 2

Spot variance path estimation and its application to high-frequency jump testing

Abstract: This chapter considers spot variance path estimation from datasets of intraday high-frequency asset prices in the presence of diurnal (or intraday seasonal) variance patterns, jumps, leverage effects and microstructure noise. We rely on parametric and nonparametric methods. The estimated spot variance path can be used to extend an existing high-frequency jump test statistic, to detect arrival times of jumps and to obtain distributional characteristics of detected jumps. The effectiveness of our approach is explored through Monte Carlo simulations. It is shown that sparse sampling for mitigating the impact of microstructure noise has an adverse effect on both spot variance estimation and jump detection. In our approach we can analyze high-frequency price observations that are contaminated with microstructure noise and rounding effects without the need for sparse sampling. An empirical illustration is presented for the intraday EUR/USD exchange rates. Our main finding is that fewer jumps are detected when sampling intervals increase.

2.1 Introduction

The increasing availability of high-frequency financial data has led to the development of new measures of variation. It is common to model financial returns as a continuous-time real-valued stochastic process such as a jump-diffusion. A possible measure of variation of such a process is quadratic variation. Realized variance is an estimate of quadratic variation and can be computed from an observed sequence of prices for some time-interval, typically

one trading day, see Andersen, Bollerslev, Diebold, and Labys (2001) and Barndorff-Nielsen and Shephard (2002). Realized variance can be employed for a wide range of applications including volatility forecasting, measuring financial risk and testing for jumps in financial price data. An up-to-date review of realized variance and its applications is presented by Andersen and Benzoni (2009). In financial market applications, quadratic variation is often decomposed into integrated variance and jump variation. Integrated variance is defined as the integral of so-called spot variance. Jump variation is defined as a sum of squared jumps and is estimated by the difference between estimates of quadratic variation and integrated variance. Estimated jump variation indicates a presence, if any, of a jump component in a given time-interval. It does not give an insight into the arrival time, size or direction of a jump. For this we require the spot variance path of high-frequency returns, see Lee and Mykland (2008). In this chapter we modify an existing estimator of integrated variance for the purpose of estimating the spot variance path and with the aim to improve testing procedures for jumps.

Most contributions for estimating spot variance are based on nonparametric methods. Under the absence of jumps and microstructure noise, estimation of spot variance is discussed, among others, in Foster and Nelson (1996), Andreou and Ghysels (2002), Alvarez, Panloup, Pontier, and Savy (2008), Fan and Wang (2008) and Kristensen (2010). Kinnebrock (2008), Mykland and Zhang (2008), Ogawa and Sanfelici (2008) and Boswijk and Zu (2009) consider estimation when high-frequency price observations are contaminated with microstructure noise. Bandi and Reno (2009) discuss estimation when price data contain either microstructure noise or jumps by making use of different consistent estimators of variation dependently on assumptions about a price process. Lee and Mykland (2008) and Boudt, Croux, and Laurent (2011) estimate spot variance in the presence of jumps, but they rule out microstructure noise. In this chapter, however, we adapt the pre-averaging approach and estimate the spot variance path when jumps and microstructure noise are both present in high-frequency price observations. In a Monte Carlo study we find that leverage effects do not have any adverse effect on the estimation performance. However, we also find that diurnal patterns can lead to systematic bias in the estimation. We therefore modify the method further to extract a periodic diurnal component in a robust manner.

Earlier theoretical and empirical contributions have focused on modifying estimation methods for integrated variance using high-frequency data to allow for jumps, leverage effects, microstructure distortions and periodic variance patterns. This work has led to mild assumptions for the spot variance process; for example, see Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006). The application of the multipower variation estimation methods for testing for jumps using high-frequency data has shown that it is necessary to allow for jumps into models for financial data. The importance of jumps is revealed by Huang and Tauchen (2005) who found that jumps account for around 7% of S&P cash index variation. Tauchen and Zhou (2010) show that for investment-grade bond spread indices, jump

variation yields more forecasting power than interest rates and other volatility measures. It illustrates the importance of developing sound statistical methods for the detection and characterization of jumps in financial returns.

Related work on high-frequency jump testing with the use of nonparametric methods is presented by Barndorff-Nielsen and Shephard (2006) where jumps are detected on the basis of a standardized difference between quadratic variation and integrated variance. A similar test for the robustness of microstructure noise is recently proposed by Podolskij and Vetter (2009a). Jiang and Oomen (2008) propose a test based on higher order return moments, that can also be employed in the presence of microstructure noise. Aït-Sahalia and Jacod (2009) use multipower variation at different sampling intervals to detect jumps. These tests are able to detect the presence of jumps in a fixed time interval. Alternatively, Andersen, Bollerslev, and Dobrev (2007) and Lee and Mykland (2008) have developed tests for finding a jump in each realized price increment. It reveals the timing as well as the distributional characteristics of realized jumps in terms of mean, variance or intensity of the jump process. Moreover, realized jumps can provide an insight into the nature of jumps for asset classes or common movements (co-jumps) between assets, see Lahaye, Laurent, and Neely (2010). Parametric methods to test for jumps are adapted by Duan and Fülöp (2007) and Bos (2008) where jumps are detected as part of the estimation method.

Boudt et al. (2011) demonstrate that pronounced intraday periodicity leads to the distorted jump inference based on the Lee and Mykland (2008) test. They modify the test to account for diurnal variance patterns. Both tests rule out microstructure noise and rely on an estimate of the spot variance path that is not robust to microstructure noise. Empirical applications of these tests are therefore based on sampling intervals that mitigate noise effects. However, sampling at high frequencies is of key importance for spot variance estimation and jump detection. Since financial time series are now widely available at high frequencies (see Table 1 of Shephard and Sheppard, 2010) sampling at low frequencies discards available information. For instance, if asset prices are observed each second, then sampling every 15 minutes discards 899 out of every 900 observations.

In this chapter we adopt a consistent estimator of integrated variance to estimate the spot variance path based on price observations which are sampled at ultra-high frequencies. The estimated spot variance path in a jump-diffusion framework with microstructure noise allows us to extend the jump test statistic of Lee and Mykland (2008). We study the impact of leverage effects, microstructure noise, rounding of prices and diurnal patterns on our proposed spot variance estimate as well as on our extended jump statistic by means of a Monte Carlo study.

We conduct an empirical study for the intraday EUR/USD exchange rates at different sampling intervals over the period of July 2 to December 31, 2007. The difference between the daily estimates of quadratic variation and integrated variance indicates the presence of jumps at the considered sampling intervals. However, when sampling intervals increase,

we are not able to detect jumps for some days even though the estimated relative jump variation has a substantial level. As shown in the simulations, the accuracy of the spot variance estimation and the power in jump detection both increase when based on finer sampling intervals.

The remainder of this chapter is organized as follows. Section 2.2 discusses the estimation of spot variance path. Section 2.3 deals with the robust estimation of the intraday return periodicity. Section 2.4 reviews the Lee-Mykland jump test and proposes an adjusted statistic. In Section 2.5 we present and discuss the results of a Monte Carlo study. Section 2.6 presents an empirical illustration, while Section 2.7 concludes.

2.2 Integrated and Spot Variance

2.2.1 General framework

For the modeling of asset prices in continuous-time, we adopt a jump-diffusion process in a similar way as Barndorff-Nielsen and Shephard (2004), Jiang and Oomen (2008), Lee and Mykland (2008) and many others. The efficient log price process X_t is assumed to follow a Brownian semi-martingale plus jumps defined as

$$\mathrm{d}X_t = \sigma_t p_t \mathrm{d}W_t + \kappa_t \mathrm{d}N_t, \qquad t \geq 0, \tag{2.1}$$

where $\sigma_t > 0$ is the stochastic part of spot volatility, $p_t > 0$ is the deterministic diurnal part of spot volatility, W_t is a standard Brownian motion, κ_t is a random variable with mean $\mu_\kappa(t)$ and variance $\sigma^2_\kappa(t)$ and N_t is a counting process that represents the number of jumps in the price path up to time t. We allow for leverage effects, that is (negative) correlation between W_t and σ_t. In this section we have $p_t = 1,\ \forall t$ in (2.1) and in Section 2.3 we take p_t as a deterministic function of time. Our aim is to estimate stochastic variance σ^2_t by means of a consistent estimator of integrated variance for a local window $[t-h,\, t]$ with $h \to 0$. We assume some degree of smoothness for the spot variance path as in Lee and Mykland (2008), Ogawa and Sanfelici (2008) and Kristensen (2010). For example, in the Monte Carlo study of Section 2.5, we use the Heston variance process driven by Brownian motion.

The price observations for (2.1) are assumed to be available at normalized equidistant times $0 = t_0 < t_1 < \ldots < t_n = 1$, where the interval $[0, 1]$ represents a trading day. Let $\Delta = t_i - t_{i-1} = n^{-1}$ be the distance between two adjacent price observations such that $\Delta \to 0$ when $n \to \infty$. The observed log price Y_{t_i} is the underlying efficient log price X_{t_i} plus error,

$$Y_{t_i} = X_{t_i} + \varepsilon_{t_i}, \tag{2.2}$$

where ε_{t_i} represents market microstructure noise with $\mathbb{E}[\varepsilon_{t_i}] = 0$, $\mathbb{E}[\varepsilon_{t_i}\varepsilon_{t_j}] = 0$ for $i \neq j$ and $\mathbb{E}[\varepsilon^2_{t_i}] = \varpi^2$. We further assume that the efficient price X_{t_i} and the microstructure noise

ε_{t_i} are independent of each other at all lags and leads, i.e. $\mathbb{E}[X_{t_i}\varepsilon_{t_j}] = 0$ for any i, j. In Section 2.4, when testing whether the price increment $Y_{t_i} - Y_{t_{i-1}}$ is contaminated by a jump, additionally it is assumed that the microstructure noise follows a Gaussian distribution. However our simulations in Section 2.5 show that this assumption can be relaxed to allow the microstructure noise be compounded by the effect of rounding prices, with little effect on the resulting jump test statistic.

If there are no jumps, this set of assumptions implies that the variance and autocovariance structure of price increments is given by

$$\mathbb{E}\left[(Y_{t_i} - Y_{t_{i-1}})(Y_{t_{i-k}} - Y_{t_{i-1-k}})\right] = \begin{cases} \int_{t_{i-1}}^{t_i} \sigma_s^2 \mathrm{d}s + 2\varpi^2, & \text{if} \quad k = 0, \\ -\varpi^2, & \text{if} \quad k = 1, \\ 0, & \text{if} \quad k > 1. \end{cases} \tag{2.3}$$

The first order autocorrelation coefficient of price increments is negative with a lower limit of $-1/2$ and all higher order autocorrelations are zero. We will use the result of (2.3) to estimate ϖ^2 from the data.

2.2.2 Integrated variance

The quadratic variation (QV) of the efficient price process (2.1) over the interval $[0, 1]$ (trading day) is defined as

$$\mathrm{QV}_{[0,1]} = \operatorname*{plim}_{\Delta \to 0} \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^2, \tag{2.4}$$

which in case of price process (2.1) can be decomposed into integrated variance (IV) and jump variation (JV)

$$\mathrm{QV}_{[0,1]} = \mathrm{IV}_{[0,1]} + \mathrm{JV}_{[0,1]}, \tag{2.5}$$

where

$$\mathrm{IV}_{[0,1]} = \int_0^1 \sigma_t^2 \mathrm{d}t \qquad \text{and} \qquad \mathrm{JV}_{[0,1]} = \sum_{j=1}^{N_1} \kappa_j^2. \tag{2.6}$$

The convergence results of the IV estimators rely on the continuous-time framework. For instance, the realized bipower variation (BPV) defined as

$$\mathrm{BPV}_{[0,1]} = \frac{\pi}{2} \sum_{i=2}^{n} |X_{t_i} - X_{t_{i-1}}||X_{t_{i-1}} - X_{t_{i-2}}|, \tag{2.7}$$

is a consistent estimator of $\mathrm{IV}_{[0,1]}$ in the absence of microstructure noise, see Barndorff-Nielsen and Shephard (2004). However, increasing the sampling frequency in the presence of microstructure noise leads to a severe upward bias in BPV, see Huang and Tauchen

(2005). For consistent estimation of IV in the presence of jumps, noise and rounding effects, one can use the pre-averaging approach. We use the pre-averaged bipower variation (PBPV) estimator, see Podolskij and Vetter (2009b), Jacod, Li, Mykland, Podolskij, and Vetter (2009a) and Podolskij and Vetter (2009a).

For applying PBPV, we select $\Theta \in (0, \infty)$ and integer k_n such that

$$k_n\sqrt{\Delta} = \Theta + o(n^{-\frac{1}{4}}), \tag{2.8}$$

and we select a weight function $q(u)$ on the $[0,1]$ interval for some variable u with $q(0) = q(1) = 0$. Podolskij and Vetter (2009a) take $q(u) = \min(u, 1-u)$ for $0 \leq u \leq 1$. The PBPV estimator is then given by

$$\text{PBPV}_{[0,1]}(l,r) = n^{\frac{l+r}{4}-1} \sum_{i=0}^{n-2k_n+1} |\bar{Y}_i|^l |\bar{Y}_{i+k_n}|^r, \tag{2.9}$$

where

$$\bar{Y}_i = \sum_{j=1}^{k_n-1} q\left(j/k_n\right)\left(Y_{t_{i+j}} - Y_{t_{i+j-1}}\right), \tag{2.10}$$

and for $l, r > 0$. The estimator (2.9) is the (l,r) order of pre-averaged bipower variation. We only consider the cases $(l,r) = (2,0)$ and $(l,r) = (1,1)$. The PBPV estimator (2.9) is biased due to microstructure noise (see Theorem 1 and 2 of Podolskij and Vetter, 2009a). The bias term depends on the microstructure noise variance ϖ^2 which we estimate on the basis of (2.3), that is

$$\widehat{\varpi}^2 = -\frac{1}{n-1}\sum_{i=2}^{n}(Y_{t_i} - Y_{t_{i-1}})(Y_{t_{i-1}} - Y_{t_{i-2}}), \tag{2.11}$$

see also Oomen (2006). The bias-corrected PBPV is then given by

$$\widehat{\text{PBPV}}_{[0,1]}(l,r) = \frac{\mu_l^{-1}\mu_r^{-1}}{\Theta\varphi_2}\text{PBPV}_{[0,1]}(l,r) - \frac{\varphi_1}{\Theta^2\varphi_2}\widehat{\varpi}^2, \tag{2.12}$$

where $\mu_p = \mathbb{E}[|z|^p]$ for $z \sim \mathcal{N}(0,1)$ such that $\mu_1 = \sqrt{2}/\sqrt{\pi}$ and $\mu_0 = \mu_2 = 1$ and where

$$\varphi_1 = \int_0^1 (q'(u))^2 \mathrm{d}u, \qquad \varphi_2 = \int_0^1 (q(u))^2 \mathrm{d}u,$$

with $q'(u)$ as the first derivative of $q(u)$ with respect to u. In case $q(u) = \min(u, 1-u)$, we have $\varphi_1 = 1$ and $\varphi_2 = \frac{1}{12}$. The choice of (l,r) and the use of non-overlapping pre-averaged statistics in (2.9) both lead to different results. While $\text{PBPV}_{[0,1]}(l,r)$ with $(l,r) =$

$(2,0)$ measures daily quadratic variation $\mathrm{QV}_{[0,1]}$, then with $(l,r)=(1,1)$ it measures daily integrated variance $\mathrm{IV}_{[0,1]}$. The relative jump variation (RJV) is estimated by

$$\widehat{\mathrm{RJV}}_{[0,1]} = 100 \times \frac{\widehat{\mathrm{PBPV}}_{[0,1]}(2,0) - \widehat{\mathrm{PBPV}}_{[0,1]}(1,1)}{\widehat{\mathrm{PBPV}}_{[0,1]}(2,0)}, \tag{2.13}$$

and measures the percentage contribution of jumps in total price variation for a trading day. Finally, Jacod et al. (2009a) show that for the special case with $q(u)=\min(u,1-u)$, $\sigma_t \equiv \sigma$ and $\mathrm{d}N_t = 0, \forall t$ in (2.1), an optimal Θ is obtained by minimizing the asymptotic variance of the estimator, and they obtain $\Theta = 4.777\,\varpi/\sigma$. We follow the authors in their suggestion to fix $\Theta = 1/3$, such that $k_n = [1/3\sqrt{n}\,]$; for instance, $k_n = 98$ when $n = 86400$ and $\Delta = 1sec$ in a $24h$ market.

2.2.3 Spot variance

The spot variance is defined as the derivative of the integrated variance

$$\sigma_t^2 = \lim_{h\to 0} \frac{\mathrm{IV}_{[t-h,t]}}{h}, \tag{2.14}$$

where $\mathrm{IV}_{[t-h,t]} = \int_{t-h}^{t} \sigma_s^2 \mathrm{d}s$, $h>0$, and can be estimated by

$$\widehat{\sigma}_t^2 = \frac{n\widehat{\mathrm{IV}}_{[t-h,t]}}{h_n}, \tag{2.15}$$

where $\widehat{\mathrm{IV}}_{[t-h,t]}$ is an unbiased and consistent estimator of $\mathrm{IV}_{[t-h,t]}$, and $h_n = hn$, with $h_n \to \infty$ satisfying $h_n/n \to 0$. The spot variance can be estimated for each observation at time t_i. In the absence of microstructure noise, the spot variance is consistently estimated by

$$\widehat{\sigma}_{t_i}^2 = \frac{\mu_1^{-2} n}{h_n} \sum_{j=i-h_n+1}^{i} |X_{t_j} - X_{t_{j-1}}||X_{t_{j-1}} - X_{t_{j-2}}|, \tag{2.16}$$

for the ith observation. However, to estimate σ_t^2 in the presence of jumps and additive noise, we adopt (2.12) to obtain

$$\widehat{\sigma}_{t_i}^2 = \frac{\mu_1^{-2} n^{1/2}}{\Theta\varphi_2 h_n} \sum_{j=i-h_n}^{i-2k_n} |\bar{Y}_j||\bar{Y}_{j+k_n}| - \frac{\varphi_1}{\Theta^2\varphi_2}\widehat{\varpi}^2, \tag{2.17}$$

for the ith observation. Since k_n is of order $n^{1/2}$, we consider $h_n = n^\beta$ with $\beta \in (1/2,\, 1]$. In case $i - h_n < 0$, we take a practical stance and use price realizations from a previous day, disregarding issues related to overnight returns. We refer to Bandi and Reno (2009) for a general theory of spot variance estimation.

The choice of the local window h for $\mathrm{IV}_{[t-h,t]}$ estimation is subject to a trade-off between bias and variance of the spot variance estimator. A small h leads to an increased variation of the estimator but it does not restrict σ_t^2 to be constant over a larger time interval. A large h assumes σ_t^2 to be constant over a large window and increases bias when spot variance is more volatile. The simulation study in Section 2.5.3 will show that different spot variance specifications and different choices of window length in (2.17) lead to different levels of estimation error.

2.3 Intraday Periodicity

2.3.1 Preliminaries

The intraday periodic patterns in return variance are mostly due to the opening, lunch-break and closing of the own market and other markets. The variance and autocorrelation functions of absolute intraday returns are typically U-shaped and persistent, see Andersen and Bollerslev (1997). We allow for diurnal patterns by decomposing spot volatility in the price process (2.1) into a stochastic variable σ_t and a smooth deterministic multiplier p_t.

2.3.2 Model with periodicity and noise

The efficient log price process (2.1) has spot volatility $\sigma_t p_t$ where p_t is higher (lower) than unity when intraday trading activity due to periodic effects increases (decreases). We further assume that

(i) $\int_0^1 p_t^2 \mathrm{d}t = 1$;

(ii) the diurnal pattern p_t is the same for each trading day;

(iii) the efficient price increment $X_{t_i} - X_{t_{i-1}}$ is observed with noise term $\varepsilon_{t_i} - \varepsilon_{t_{i-1}}$ that has mean zero and variance $2\varpi^2 p_{t_i}^2$;

In practice, assumption (i) is sufficient to ensure that the integrated variance $\mathrm{IV}_{[0,1]}$ estimate for an entire trading day is not affected by intraday periodicity p_t, see also the discussion in Barndorff-Nielsen and Shephard (2004). However, in case we estimate $\sigma_t p_t$ using observations from an interval smaller than one trading day, diurnal patterns do matter. Assumption (ii) can easily be relaxed to allow the diurnal pattern to depend on the day-of-the-week. This possibility will be used in the application. Assumption (iii) is motivated by Kalnina and Linton (2008) who have considered equity data and estimated U-shaped intraday variance patterns in microstructure noise. However, it does imply that

the noise-to-signal ratio

$$\mathrm{NSR}_{t_i} = \frac{2\varpi^2 p_{t_i}^2}{\sigma_{t_i}^2 p_{t_i}^2 \Delta + 2\varpi^2 p_{t_i}^2} = \frac{2\varpi^2}{\sigma_{t_i}^2 \Delta + 2\varpi^2}, \tag{2.18}$$

is not subject to a diurnal shape. Otherwise, the constant Θ in the pre-averaged based estimator (2.8) would be a function of the diurnal shape p_t. The issue of periodic behavior of Θ is not considered here. The diurnal effect is therefore a strict multiplicative effect and it does not affect the noise-to-signal ratio.

2.3.3 Periodicity extraction

Assumption (ii) implies that the estimation of the diurnal effect can be based on financial returns from multiple trading days. For this purpose, we denote the observed price, spot variance, microstructure noise variance and pre-averaged bipower estimator (2.12) at time t_i on day d by

$$Y_{d+t_i}, \qquad \sigma^2_{d+t_i}, \qquad \varpi^2_d, \qquad \text{and} \qquad \widehat{\mathrm{PBPV}}_d(l,r),$$

respectively, for days $d = 1, \ldots, D$. We follow Boudt et al. (2011) in benchmark estimation of the periodic component p_{t_i} on the basis of cross-sectional variances. In case of no jumps in the price process (2.1), the mean and variance of the price increment

$$Y_{d+t_i} - Y_{d+t_{i-1}} = X_{d+t_i} - X_{d+t_{i-1}} + \varepsilon_{d+t_i} - \varepsilon_{d+t_{i-1}}, \tag{2.19}$$

are given by

$$\mathbb{E}\left[Y_{d+t_i} - Y_{d+t_{i-1}}\right] = 0 \quad \text{and} \quad \mathbb{V}ar\left[Y_{d+t_i} - Y_{d+t_{i-1}}\right] = \sigma^2_{d+t_i} p^2_{t_i} \Delta + 2\varpi^2_d p^2_{t_i}. \tag{2.20}$$

Given assumption (ii), price increments at time t_i for all D days have a common deterministic multiplier p_{t_i}. When all price increments are standardized based on $\sigma^2_{d+t_i}\Delta + 2\varpi^2_d$, then the variance of the standardized increments is equal to $p^2_{t_i}$. The cross-section over days of the standardized increments at time t_i is used to estimate $p^2_{t_i}$ using all D days.

The estimation of $p^2_{t_i}$ requires estimates for $\sigma^2_{d+t_i}$ and ϖ^2_d. We treat $\sigma^2_{d+t_i} \equiv \sigma^2_d$ as constant over the dth day and $\widehat{\mathrm{PBPV}}_{[0,1]}(1,1)$ of (2.12) provides its estimate. We take (2.11) as an estimate for ϖ^2_d. To satisfy assumption (i), we normalize the diurnal estimator for $p^2_{t_i}$ by

$$\widehat{p}^2_{t_i} = \frac{n \sum_{d=1}^{D} R^2_{d+t_i}}{\sum_{j=1}^{n}\sum_{d=1}^{D} R^2_{d+t_j}}, \tag{2.21}$$

where

$$R^2_{d+t_i} = \frac{\left(Y_{d+t_i} - Y_{d+t_{i-1}}\right)^2}{n^{-1}\widehat{\mathrm{PBPV}}_{[0,1]}(1,1) + 2\widehat{\varpi}^2_d}, \tag{2.22}$$

for $i = 1, \ldots, n$. Our approach differs from Andersen and Bollerslev (1998b) in which σ_d^2 is estimated on the basis of a generalized autoregressive conditional heteroskedasticity model for daily returns and hence does not use intra-daily data for the estimation of the daily variance.

2.3.4 Robust estimation of intraday periodicity

In case jumps are present in the price process (2.1), we need to modify estimator (2.21) to keep it smooth. Since theory provides no specific guidelines regarding the shape of intraday periodicity, we propose to locally smooth the cross-sectional variances (2.21) via the weighted average

$$\widehat{\bar{p}}_{t_i}^2 = \sum_{j=1}^{n} \omega_{ij} \widehat{p}_{t_j}^2, \tag{2.23}$$

where ω_{ij} is the normalized weight such that $\sum_{j=1}^{n} \omega_{ij} = 1$, for $i = 1, \ldots, n$. The weights are obtained from a two-sided exponentially weighted moving average (EWMA) smoothing scheme and depend on a discounting factor. The EWMA weighting scheme can be implicitly obtained from a local level model in levels, as discussed in Durbin and Koopman (2001, §3.4). The local level model for the periodic estimates $\widehat{p}_{t_j}^2$ is given by

$$\widehat{p}_{t_i}^2 = \bar{p}_{t_i}^2 + \zeta_{t_i}, \tag{2.24}$$

$$\bar{p}_{t_{i+1}}^2 = \bar{p}_{t_i}^2 + \eta_{t_i}, \tag{2.25}$$

where we treat $\widehat{p}_{t_i}^2$ as the observation and $\bar{p}_{t_i}^2$ as the unknown signal that needs to be estimated. The disturbances ζ_{t_i} and η_{t_i} are mutually and serially independent with mean zero and variances σ_ζ^2 and σ_η^2, respectively. The signal is modeled as a random walk process with Gaussian disturbances (or increments) η_{t_i}. By applying the Kalman filter and smoother to the local level model (2.24)-(2.25) we obtain the estimate for $\bar{p}_{t_i}^2$ that is computed as (2.23) with the discount factor depending on the variance ratio $\sigma_\eta^2/\sigma_\zeta^2$. The variances can be estimated by the method of maximum likelihood where the loglikelihood function of the local level model is obtained from the Kalman filter; see Durbin and Koopman (2001, Ch. 2).

We expect that jumps in the return series will cause outliers in the "observations" $\widehat{p}_{t_i}^2$. To obtain robust estimates for $\bar{p}_{t_i}^2$, we let the disturbances ζ_{t_i} to come from a Student's t distribution. In Harvey and Koopman (2000) the effect of the Student's t distribution for ζ_{t_i} on outlying observations $\widehat{p}_{t_j}^2$ is illustrated in detail. The outlying observations receive weights ω_{ij} which are close to zero, depending on the severeness of the outlier in relation to the degrees of freedom of the t distribution. The jumps do not influence the smoothed

periodicity estimate as a result.[1] We emphasize that the local level model for $\widehat{p}^2_{t_i}$ does not have any implications for our theoretical framework in Section 2; it only facilitates the construction of robust EWMA weights in (2.23). The estimation of parameters (including the degrees of freedom) and the signal $\bar{p}^2_{t_i}$ of the Student's t local level model relies on computationally efficient Monte Carlo methods and are based on importance sampling techniques; see Durbin and Koopman (2001) for a detailed discussion. The computations are carried out in `Ox` (Doornik, 2009) using `SsfPack` of Doornik, Koopman, and Shephard (1999).

2.4 Intraday Jump Testing

2.4.1 Lee and Mykland test

Consider the price process (2.1) with $p_t = 1$ and assume that a realization X_{t_i} from (2.1) can be observed without noise. The jump test statistic of Lee and Mykland (2008) standardizes the increment $X_{t_i} - X_{t_{i-1}}$ by a local measure of variation for the continuous part of the price process. In case a jump occurs in the $(t_{i-1}, t_i]$ interval, the realized price increment $X_{t_i} - X_{t_{i-1}}$ is obviously larger than we can expect from the continuous part of the price process. The Lee and Mykland jump test statistic (LMJ) exploits this notion and proposes a nonparametric statistic to test the hypothesis whether a jump has occurred in the interval $(t_{i-1}, t_i]$. The LMJ test $\mathcal{L}_{\text{LMJ}}$ is based on the asymptotic distribution of the maximums of absolute values of standard normal variables,

$$\mathcal{L}_{\text{LMJ}} = \frac{\max_{i\in[1,n]}|\mathcal{L}(t_i)| - C_n}{S_n}, \tag{2.26}$$

where the constants C_n and S_n are given by

$$S_n = (2\log n)^{-\frac{1}{2}} \qquad \text{and} \qquad C_n = S_n^{-1} - \frac{1}{2}(\log\pi + \log\log n)S_n, \tag{2.27}$$

and where

$$\mathcal{L}(t_i) = \frac{X_{t_i} - X_{t_{i-1}}}{\widehat{\sigma}_{t_i}\sqrt{\Delta}}, \tag{2.28}$$

is standard normally distributed under the null hypothesis of no jumps. Lee and Mykland (2008) compute the spot variance estimate $\widehat{\sigma}^2_{t_i}$ by the locally averaged bipower variation (2.16) over the window $[t_{i-h_n}, t_{i-1}]$ with sufficiently large h_n, say $h_n = 270$ for $\Delta = 5min$. The test statistic $\mathcal{L}_{\text{LMJ}}$ in (2.26) has a standard Gumbel distribution. The null hypothesis of no jump in $(t_{i-1}, t_i]$ is rejected when $|\mathcal{L}(t_i)| > S_n\alpha^* + C_n$ at the significance level $\alpha\%$ with $\alpha^* = -\log(-\log(1-\alpha))$.

[1] In the simulation exercise this effect is most clearly seen from Figure 2.1 in panels *iii-b)* to *v-b)*, where the contaminated periodicity in period 15:00-15:01 is given an effective weight of zero.

The arrival of a jump of size κ^* in the interval $(t_{i-1}, t_i]$ leads to an increase of $X_{t_i} - X_{t_{i-1}}$ by κ^* and thus $\mathcal{L}(t_i)$ in (2.28) is not standard normally distributed. The jump size κ^* in $\mathcal{L}(t_i)$ is divided by $\widehat{\sigma}_{t_i}\sqrt{\Delta}$ and $\mathcal{L}(t_i) \to \infty$ when $\Delta \to 0$, so statistic becomes sufficiently large to detect the jump arrival as $\Delta \to 0$, implying that more frequent sampling increases the power of the test (see Theorem 2 of Lee and Mykland, 2008).

2.4.2 A jump test correction for periodicity and microstructure noise

When the locally averaged bipower variation estimator is used for spot variance in the presence of microstructure noise, the statistic $\mathcal{L}(t_i)$ in (2.28) is not standard normal under the null hypothesis of no jumps. Therefore, the LMJ test statistic (2.26) is not valid when microstructure noise is present. Also, when the intraday return variation is subject to periodicity, Andersen et al. (2007) and Boudt et al. (2011) argue that spurious jumps can be detected. Here we correct the LMJ test for both periodicity and microstructure noise.

Given the assumptions on the microstructure noise in Section 2.2.1, assumptions (i)-(iii) of Section 2.3.2 and under the null hypothesis of no jumps in the price process (2.1), the realized price increment (2.19) is normally distributed with mean and variance given by (2.20). We adopt these results to modify the LMJ test statistic (2.26) by replacing $\mathcal{L}(t_i)$ with $\widetilde{\mathcal{L}}(t_i)$ as given by

$$\widetilde{\mathcal{L}}(t_i) = \frac{Y_{t_i} - Y_{t_{i-1}}}{\widehat{s}_{t_i}}, \tag{2.29}$$

where $\widehat{s}^2_{t_i} = \widehat{\sigma}^2_{t_i}\widehat{\overline{p}}^2_{t_i}\Delta + 2\widehat{\varpi}^2\widehat{\overline{p}}^2_{t_i}$, with the periodic multiplier $p^2_{t_i}$ estimated as described in Section 2.3.4 and where $\widehat{\sigma}^2_{t_i}$ and $\widehat{\varpi}^2$ are given by (2.17) and (2.11), respectively. The variable $\widetilde{\mathcal{L}}(t_i)$ is standard normal and our corrected test statistic $\widetilde{\mathcal{L}}_{\text{LMJ}}$ (cf. (2.26)) remains to have a Gumbel distribution under the null hypothesis of no jumps.

The arrival of a jump of size κ^* in the interval $(t_{i-1}, t_i]$ leads to an increase of $Y_{t_i} - Y_{t_{i-1}}$ by κ^* and thus $\widetilde{\mathcal{L}}(t_i)$ in (2.29) is not standard normally distributed. The jump size κ^* in $\widetilde{\mathcal{L}}(t_i)$ is divided by $\widehat{s}_{t_i}$ and the discontinuous part is now proportional to $\kappa^*/\sqrt{2}\varpi p_{t_i}$ as $\Delta \to 0$. Hence, jump detection in the presence of microstructure noise depends on the magnitude of price contamination. When the standard deviation of microstructure noise increases, the statistic $\widetilde{\mathcal{L}}(t_i)$ decreases, with the implication that small jumps are more difficult to detect.

We follow Hansen and Lunde (2006) and the discussion by Diebold (2006) who argue that the assumption of normality for the microstructure noise is not plausible when the price discreteness is taken into account. When the price increments $Y_{t_i} - Y_{t_{i-1}}$ are living on a grid defined by the tick-size, the distributions of $\widetilde{\mathcal{L}}(t_i)$ and related $\widetilde{\mathcal{L}}_{\text{LMJ}}$ can be distorted. When the distributions of the one-period price increments give support to multiple ticks, the density of the test statistics $\widetilde{\mathcal{L}}_{\text{LMJ}}$ can be approximated accurately by the Gumbel density;

we provide simulation evidence in Section 2.5.4. Multiple ticks are typically obtained when the volatility is sufficiently high and/or the rounding effect is small.

2.5 Monte Carlo Studies

In this section we conduct a set of Monte Carlo experiments to study the finite sample properties of the intraday periodicity and the spot variance path estimation procedures as discussed in the previous sections. Furthermore we study the small-sample performance of our intraday jump testing procedure. We first present the design of the Monte Carlo study.

2.5.1 Design of Monte Carlo study

The price series are generated by (2.1) which depends on assumptions for spot volatility σ_t, periodic multiplier p_t and random jump process κ_t. We adopt the mean-reverting Heston (1993) process for the spot variance σ_t^2 that is given by

$$\mathrm{d}\sigma_t^2 = \psi(\theta - \sigma_t^2)\mathrm{d}t + \gamma\sigma_t \mathrm{d}W_t^\sigma, \tag{2.30}$$

where coefficient $\psi > 0$ determines the rate at which the variance σ_t^2 reverts to the long-run variance $\theta > 0$, coefficient γ determines the variation of the spot variance path and W_t^σ is a standard Brownian motion. The Brownian motions W_t in (2.1) and W_t^σ are possibly contemporaneously correlated, that is $\mathbb{E}[\mathrm{d}W_t \mathrm{d}W_t^\sigma] = \rho \mathrm{d}t$ with $\rho \in (-1, 0]$. In case $\rho \neq 0$, the variance process is subject to a leverage effect. The multiplicative periodic factor p_t in (2.1) is assumed to be a smooth function. The jump component κ_t in (2.1) is given by the product $\bar{\kappa}\sqrt{\theta} p_t U_t$ where $\bar{\kappa}$ determines the size of the jump relative to the long-run volatility and periodicity, and where U_t is a random variable which takes the values -1 and 1 with equal probability. To produce the process N_t in (2.1) we allocate via the homogeneous Poisson process at least one jump over a trading day.

We simulate sample paths for the price process with $X_0 = 0$ using the Euler discretization method at a time interval $t_{i+1} - t_i = 86400^{-1}$, corresponding to the one second frequency in a 24h market. Then we sample observations from the price path at different frequencies Δ; by increasing Δ the sampling becomes more sparse, which mitigates microstructure effects. The price increment $X_{t_i} - X_{t_{i-1}}$ is contaminated by microstructure noise as implied by assumption (iii) of Section 2.3.2.

Our Monte Carlo design is based on the following settings:

- For each experiment, we sample $M = 2500$ realizations of the prices process. Each realization is a time series of prices for a single 24h trading day. The length of the time series depends on sampling frequency $\Delta \in \{1sec, 5sec, 30sec, 1min, 5min\}$.

- The stochastic variance specification (2.30) is adopted from Heston (1993) and is given by the two variants

  i) Heston-A: $\psi = 1$, $\theta = 0.04$, $\gamma = 0.15$;

  ii) Heston-B: $\psi = 8$, $\theta = 0.04$, $\gamma = 0.4$.

  where both variants satisfy the condition $2\psi\theta \geq \gamma^2$ such that the variance is strictly positive. The Heston-A specification has the smallest mean-revision parameter and the smallest value for γ. In both specifications we have $\theta = 0.04$ that corresponds to a long-run volatility level of 20%.

- The multiplicative periodic diurnal factor is given by $p_{t_i}^2 = 1 - \phi\cos(2\pi i/n)$ with $\phi \in [0, 1)$. For $\phi = 0$, we have no diurnal shape in the return variance and in the microstructure noise variance.

- The jump size is taken from the set $\bar{\kappa} \in \{0,\ 0.05,\ 0.15,\ 0.25,\ 0.5\}$. For these values of $\bar{\kappa}$, the mean estimates of relative jump variation in (2.13) at $\Delta = 1sec$ have values $\widehat{\mathrm{RJV}}_{[0,1]} \in \{0.22\%,\ 0.34\%,\ 2.24\%,\ 6.63\%,\ 23.26\%\}$ and $\widehat{\mathrm{RJV}}_{[0,1]} \in \{0.22\%,\ 0.31\%,\ 2.21\%,\ 6.59\%,\ 23.19\%\}$ with $\varpi = 0.001$ and $\varpi = 0.004$ respectively. When the data is sampled more sparsely and/or microstructure noise increases, the impact of jumps decreases. In case $\bar{\kappa} \neq 0$, we enforce at least one jump in a trading day.

- The microstructure noise standard deviation is taken from the set $\varpi \in \{0,\ 0.001,\ 0.004\}$. For positive values of ϖ, the ratios of microstructure noise variance and integrated variance correspond to values found in equity data by Hansen and Lunde (2006, Table 3).

- When a Monte Carlo experiment requires the estimation of periodic factors, we repeat the simulations for $D = 100$ trading days. For all other experiments we set $D = 1$.

- The effect of rounding prices, on estimation of spot variance and on simulated jump test statistics, is studied as well. We consider rounding prices to the nearest multiple of $g \in \{0.001, 0.005, 0.01\}$, for a light, moderate and severe level of rounding.

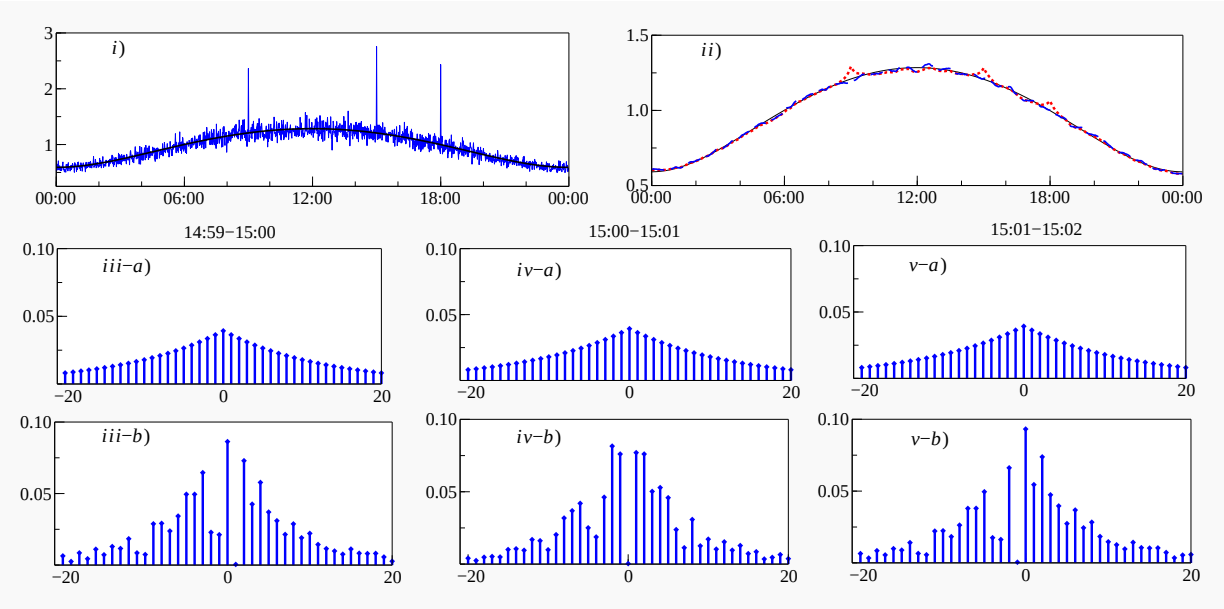
Each experiment depends on the choice of parameters: ϕ (periodicity), $\bar{\kappa}$ (jump size), ϖ (standard deviation of microstructure noise) and ρ (leverage effect). Their subsequent values are specified for each experiment and reported in the discussions, figures and tables.

2.5.2 Periodic patterns in spot variance

We first investigate whether the cross-sectional estimation of the periodic pattern (2.21) in the spot variance is robust to jumps. To measure the impact of jumps and its sensitivity to

different weight adjustments in (2.23), we carry out the following experiment. Given the features of the empirical data in the next section, we consider jumps arriving at fixed time arrivals (news announcements at 9:00-9.01, 15.00-15.01 and 18.00-18.01) and at random time arrivals. The jumps themselves are drawn with $\bar{\kappa} = 0.15$. The remaining Monte Carlo parameters are given by $\phi = 0.65$, $\varpi = 0.001$ and $\rho = 0$.

Figure 2.1: *Periodicity extraction*



Note: Heston-A specification with $\phi = 0.65$, $\bar{\kappa} = 0.15$, $\varpi = 0.001$, $\rho = 0$, $\Delta = 1min$ and $D = 100$. Panel *i)* presents square root of cross-sectional variances (dotted line) and true periodic factor (solid line); *ii)* presents estimates from OWM-N (dotted line), OWM-t (dashed line) and true periodic factor (solid line); panels *iii)*, *iv)* and *v)* present ω_{ij} in (2.23) for 20 lags/leads for OWM-N *iii-a)* − *v-a)*, and OWM-t, *iii-b)* − *v-b)*, in time interval 14:59-15:00, 15:00-15:01 and 15:01-15:02 respectively.

Figure 2.1 presents the estimated periodic pattern. The benchmark estimate of the diurnal shape from (2.21) shown in panel *i)* is affected by jumps. The scheduled news announcements have an adverse effect. The smoothing method in (2.23) is based on the model (2.24)-(2.25) with Gaussian or Student's t disturbances in (2.24). The observation based weighting method (OWM) in (2.23) with Gaussian (Student's t) observation disturbances is denoted as OWM-N (OWM-t). Both estimates are presented in panel *ii)* of Figure 2.1 and are more smooth relative to benchmark estimate. However, the symmetric weights ω_{ij} implied by the model with Gaussian observation disturbances do not adjust sufficiently at the jump location and its neighboring time intervals; see panels *iii-a)* to *v-a)* of Figure 2.1. The weights implied by Student's t observation disturbances are small or zero at jump locations; see panels *iii-b)* to *v-b)* of Figure 2.1. Also, neighboring weights are not affected by the actual jumps. We can conclude that our last and preferred extraction procedure is robust to jumps.

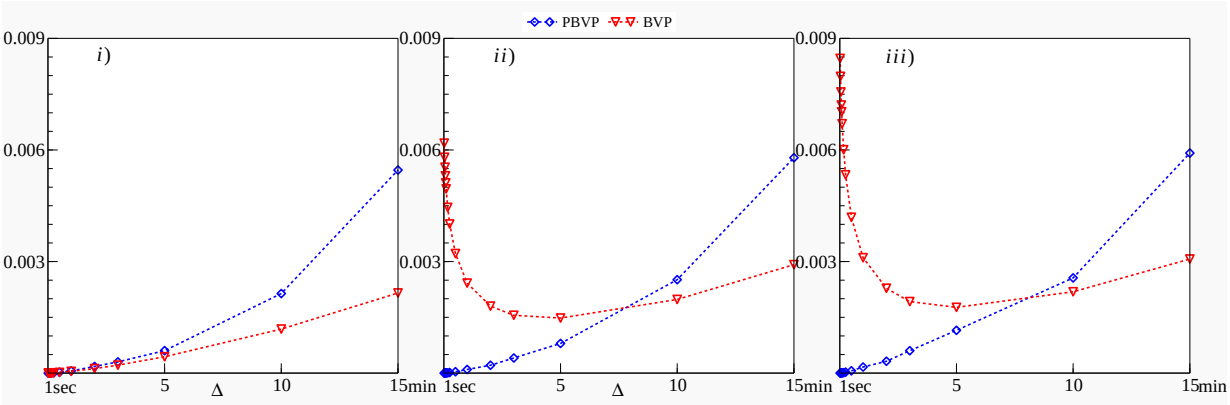
2.5.3 Spot volatility estimation

This section considers the spot volatility estimation with regards to microstructure noise, periodic pattern, window width and leverage effect.

Presence of microstructure noise and rounding

We focus on the effect of microstructure noise and compare performance of σ_t^2 estimator based on bipower variation (2.16) and pre-averaged bipower variation (2.17). The Monte Carlo parameters are given by $\phi = 0, \rho = 0$ (no periodicity and no leverage), $\varpi \in \{0, 0.001\}$ and rounding effect $g \in \{0, 0.005\}$.

Figure 2.2: *Comparison of spot variance estimators*



Note: Root mean squared error (RMSE) of spot variance against sampling frequency. Panel *i)* with no microstructure noise $\varpi = 0$; *ii)* with microstructure noise $\varpi = 0.001$, *iii)* with microstructure noise $\varpi = 0.001$ and rounding prices to the nearest multiple of $g = 0.005$. Triangles correspond to spot variance estimator based on bipower variation (2.16), diamonds to pre-averaged based estimator (2.17). Heston-B specification with $h_n = n^{0.8}$, $\phi = 0$, $\rho = 0$, $\bar{\kappa} = 0.25$.

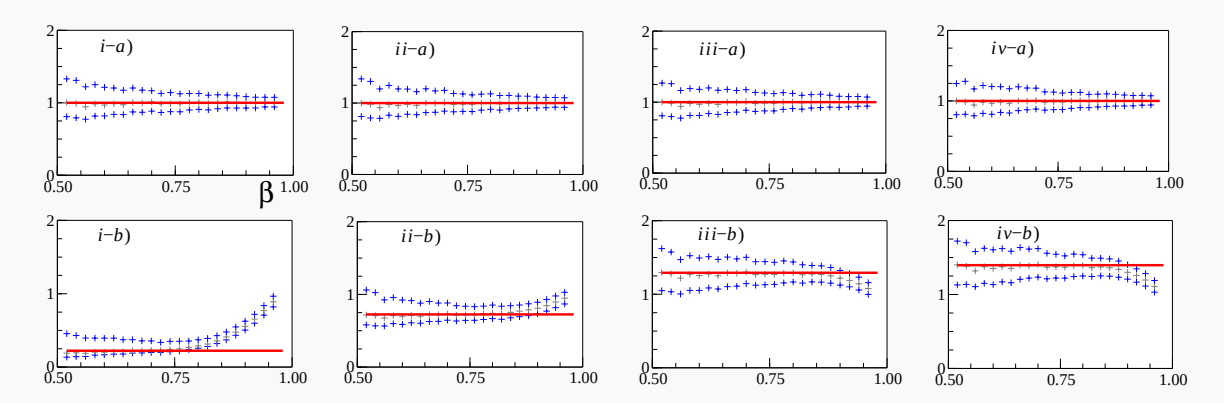
Figure 2.2 presents the root mean squared error (RMSE) of the estimated spot variance against the sampling frequency. Without microstructure noise or rounding, the best estimation performance for both estimators is obtained at the finest frequency as shown in panel *i)*. The bipower variation based spot variance estimator (2.16) has lower RMSE than the pre-averaged based estimator (2.17) when the sampling frequency increases. With microstructure noise, the (2.16) estimator is not optimal for the finest sampling frequency; see panel *ii)* of Figure 2.2. The RMSE dramatically increases when $\Delta \to 0$ for the bipower variation estimator (2.16). On the other hand, the pre-averaged based estimator (2.17) has the lowest RMSE when data is sampled at the finest frequency. It shows that the correction for bias due to microstructure noise is effective. There is no need to search for an optimal sampling frequency as in the case of bipower variation. Finally, we consider a combination of additive noise and moderate rounding in panel *iii)*. In general, the pre-averaging step regularizes the price process and reduces the effects of price discreteness. It

is seen that rounding-off prices, as an additional source of frictions, does not impact the pre-averaged based estimator of spot variance, while the upward bias of bipower variation increases further when $\Delta \to 0$.

Effect of periodic patterns

Here we show the importance of having separate estimation procedures for the spot variance components σ_t and p_t in (2.1). We focus on the impact of periodicity on the joint estimation of $\sigma_t p_t$ based on pre-averaged estimator. We consider jumps and noise with the Monte Carlo parameters given by $\varpi = 0.001$, $\rho = 0$ and $\bar{\kappa} = 0.25$.

Figure 2.3: *Bias due to intraday periodicity*



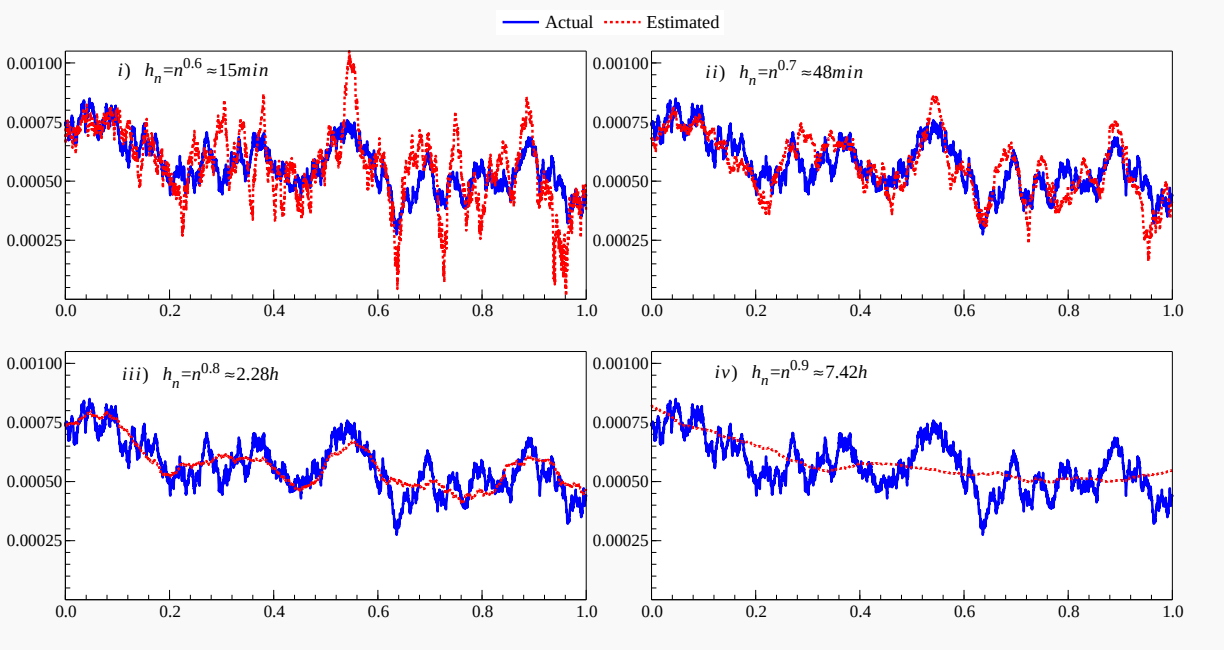
Note: Estimates of spot volatility as a function of β; $h_n = n^\beta$. Crosses correspond to 2.5%, 50% and 97.5% quantiles based on $M = 2500$, solid line is the true value. Panels correspond to four different points of time: *i)* 00:00; *ii)* 03:00; *iii)* 09:00; *iv)* 12:00. Top panels with $\phi = 0$ (no diurnal pattern) and bottom panels with $\phi = 0.95$ (diurnal pattern). For both, $\rho = 0$, $\bar{\kappa} = 0.25$, $\varpi = 0.001$ and $\Delta = 1sec$.

Figure 2.3 presents estimates of $\sigma_t p_t$ obtained by using different window widths determined by β, i.e. $h_n = n^\beta$ for $\beta \in (1/2, 1]$. Four different intraday points of time are shown. The top panels of Figure 2.3 correspond to $\phi = 0$ (no periodicity) and the bottom panels show impact of diurnal shape with $\phi = 0.95$. Even though the periodic variance is smooth and data is sampled at the highest frequency $\Delta = 1sec$, the presence of diurnal pattern induces a systematic bias which increases when the width of the local window gets larger. Periodic patterns have therefore an adverse effect on the joint estimation of the spot variance. The systematic bias increases when the data is sampled more sparsely and with more pronounced periodic patterns.

Selection of window width

Next we turn our attention to the estimation of the stochastic component in spot volatility and the choice of the window length in (2.17). We consider $h_n = n^\beta$ with $\beta \in \{0.6, 0.7, 0.8, 0.9\}$.

Figure 2.4: *Choosing the window width*



Note: An example of actual (solid line) vs. estimated (dashed line) spot volatility path based on $h_n = n^{\beta}$. Heston-A specification with $\phi = 0$, $\rho = 0$, $\bar{\kappa} = 0.25$, $\varpi = 0.001$, $\Delta = 1sec$

Figure 2.4 presents estimated paths over a trading day for the Heston-A variance specification. Estimation based on the small window $n^{0.6}$ leads to high variation of the spot variance estimator; see panel *i)*. When the long window $n^{0.9}$ is taken, a highly smooth estimate is obtained, see panel *iv)* of Figure 2.4.

Table 2.1 presents simulation results for the selection of the optimal bandwidth for Heston-A and Heston-B specifications. We obtain the following findings. For both specifications, the mean squared error (MSE) increases when the sampling is more sparse. For all sampling frequencies, the measures of fit for Heston-A are more satisfactory than the ones for Heston-B. The lowest MSEs are obtained for $h_n = n^{0.8}$ (Heston-A) and $h_n = n^{0.7}$ (Heston-B). This result is robust up to the sampling frequency $\Delta = 5min$.

When we allow for a leverage effect with $\rho = -0.75$, we do not find any adverse effect on the estimation performance; see the lowest panels of Table 2.1, where we compare the resulting MSEs relative to the no leverage case with $\rho = 0$. The obtained MSEs are very much the same with ratios oscillating around unity.

2.5.4 Intraday jump testing

In this section we investigate the distribution of the test statistic under the null of no jumps in the price process. Also, effective power and size are studied, and how these are

Table 2.1: *Bandwidth selection*

	$\Delta = 1sec$	$\Delta = 5sec$	$\Delta = 30sec$	$\Delta = 1min$	$\Delta = 5min$
			Heston-A		
$\rho = 0$					
$\beta = 0.6$	2.066	1.614	2.787	2.046	5.695
$\beta = 0.7$	1.186	1.062	1.522	1.128	1.897
$\beta = 0.8$	**1.000**[0.001]	**1.000**[0.011]	**1.000**[0.113]	**1.000**[0.283]	1.328
$\beta = 0.9$	1.778	1.284	1.161	1.098	**1.000**[2.403]
$\rho = -0.75$					
$\beta = 0.6$	2.107	1.600	2.823	2.052	5.552
$\beta = 0.7$	1.207	1.051	1.535	1.128	1.846
$\beta = 0.8$	**1.000**[0.001]	**1.000**[0.011]	**1.000**[0.112]	**1.000**[0.280]	1.333
$\beta = 0.9$	1.788	1.300	1.174	1.005	**1.000**[2.407]
$\rho = -0.75/\rho = 0$					
$\beta = 0.6$	1.019	0.982	1.001	0.993	0.977
$\beta = 0.7$	1.017	0.988	0.997	0.990	0.975
$\beta = 0.8$	0.999	0.991	0.988	0.990	1.006
$\beta = 0.9$	1.004	1.003	1.001	0.996	1.002
			Heston-B		
$\rho = 0$					
$\beta = 0.6$	1.134	1.029	1.382	1.241	2.592
$\beta = 0.7$	**1.000**[0.003]	**1.000**[0.021]	**1.000**[0.256]	**1.000**[0.546]	1.119
$\beta = 0.8$	1.777	1.644	1.177	1.311	**1.000**[5.743]
$\beta = 0.9$	4.376	3.125	1.834	1.891	1.107
$\rho = -0.75$					
$\beta = 0.6$	1.139	1.025	1.389	1.237	2.588
$\beta = 0.7$	**1.000**[0.003]	**1.000**[0.021]	**1.000**[0.259]	**1.000**[0.550]	1.120
$\beta = 0.8$	1.768	1.649	1.176	1.320	**1.000**[5.732]
$\beta = 0.9$	4.371	3.122	1.849	1.903	1.122
$\rho = -0.75/\rho = 0$					
$\beta = 0.6$	1.008	0.997	1.009	1.005	0.997
$\beta = 0.7$	1.004	1.000	0.997	1.009	0.999
$\beta = 0.8$	0.998	1.004	0.996	1.016	0.998
$\beta = 0.9$	1.002	0.999	1.004	1.015	1.011

Note: The entries present the mean-squared errors (MSE) relative to the smallest one (between brackets $\times 10^6$) within the model setup, e.g. $\{\Delta = 1sec, \rho = 0\}$. Panel $\rho = -0.75/\rho = 0$ compares MSEs with and without leverage effect. It holds, $\phi = 0$, $\varpi = 0.001$, $\bar{\kappa} = 0.25$. Estimation done over a local window $h_n = n^\beta$ with $\beta = \{0.6; 0.7; 0.8; 0.9\}$.

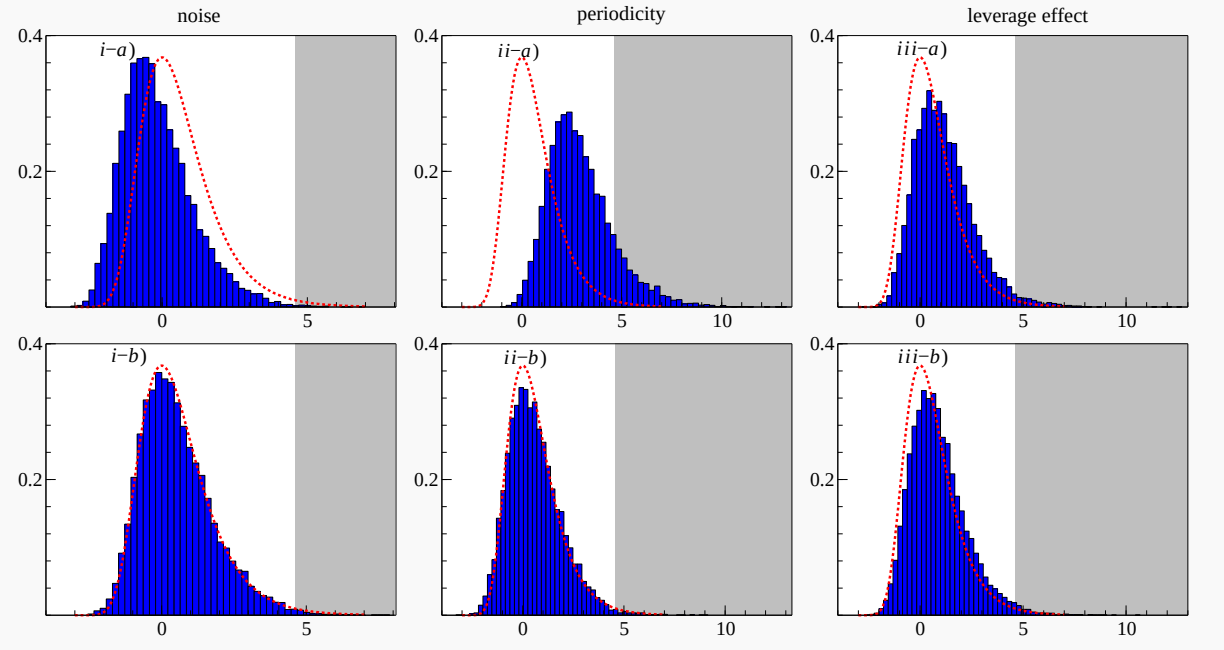
influenced by possible rounding of the prices.

Null distribution of the jump test statistic without rounding

We study the distributions of the jump test statistic based on empirical densities of the maximums of the test statistic under the absence of jumps and compare those to the theoretical standard Gumbel distribution. We compute the empirical density of test statistic (2.26) based on the original $\mathcal{L}(t_i)$ in (2.28), and the extended test statistic based on $\widetilde{\mathcal{L}}(t_i)$ in (2.29).

Figure 2.5 presents empirical densities of simulated test statistics. To see the impact of microstructure noise, periodicity and leverage effect, we study the empirical distributions under the null separately for each of these features shown in panels *i)-iii)*.

Figure 2.5: *Simulated test densities*



Note: Theoretical density of the jump test statistic (dotted line) versus empirical density under the null of no jumps. Panels *a)* for LMJ test (2.26) based on original $\mathcal{L}(t_i)$ in (2.28); panels *b)* for extended LMJ test based on $\widetilde{\mathcal{L}}(t_i)$ in (2.29). Heston-A specification with $\Delta = 1sec$ and $M = 2500$. Panels *i)* present effect of noise with $\varpi = 0.004$; panels *ii*) of periodicity with $\phi = 0.95$; panels *iii)* present effect of leverage with $\rho = -0.75$. Grey area is the rejection region at the $\alpha = 1\%$ significance level; $\alpha^* = 4.6001$.

The left-hand panels show the impact of microstructure noise with $\varpi = 0.004$. The empirical distribution of the original LMJ test based on $\mathcal{L}(t_i)$ is affected, as the mass of the density shifts to the left as seen in *i-a)* relative to the theoretical density. Thus, instead of bipower variation (2.16) we apply pre-averaged bipower variation (2.17). Further, the variance of microstructure noise has to be accounted for, because when $\Delta \to 0$, then NSR in (2.18) approaches unity and a larger proportion of return variation is due to microstructure effects. The empirical density of the extended LMJ test $\widetilde{\mathcal{L}}_{\text{LMJ}}$ based on $\widetilde{\mathcal{L}}(t_i)$ coincides with the theoretical Gumbel density; see panel *i-b)*.

The impact of the diurnal pattern with $\phi = 0.95$ is shown in the middle panels. Intraday periodicity has an adverse effect on the LMJ test as seen in *ii-a)*, where it overrejects the null of no jump. This effect was noticed before by Boudt et al. (2011). Pronounced diurnal patterns lead to systematic bias in the spot variance estimator, unless a very small local window is selected at the cost of lack of robustness against jumps and inflated variation of the estimator; see the discussion in Section 2.5.3. We estimate the spot variance components separately and panel *ii-b)* shows the appropriate empirical density of extended test based on $\widetilde{\mathcal{L}}(t_i)$.

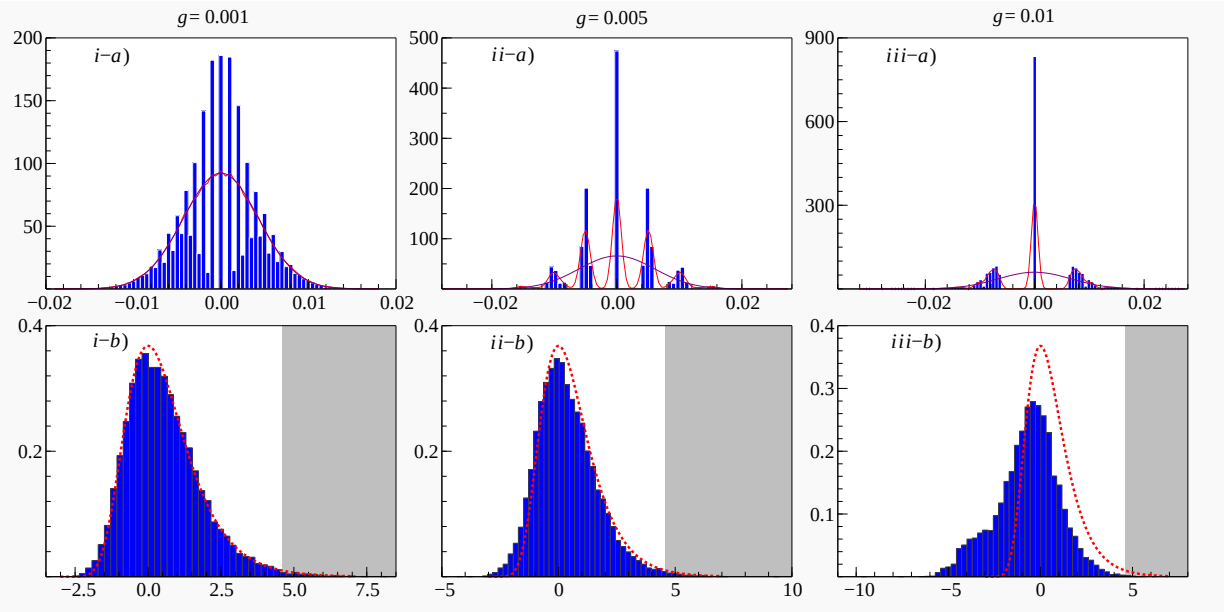
We allow for the leverage effect with $\rho = -0.75$ and present the resulting empirical

densities in the right-hand panels. The negative correlation in price-variance innovations slightly shifts the mass of the empirical densities to the right. It is more apparent in the original LMJ test based on $\mathcal{L}(t_i)$ as seen in panel *iii-a)*. Both tests can possibly overreject the null of no jump due to the leverage.

Effect of rounding on the jump test statistic

When prices live on a grid defined by the tick-size g, the distribution of the test statistic $\widetilde{\mathcal{L}}_{\text{LMJ}}$ can be distorted.

Figure 2.6: *Effect of price discreteness*



Note: Theoretical density of the extended jump test statistic (2.29) versus empirical density under the null of no jump. The panels show the impact of additive microstructure noise $\varpi = 0.004$ combined with rounding prices to levels of multiples of $g = 0.001$ (left panels), $g = 0.005$ (middle panels) and $g = 0.01$ (right panels). The top panels present an example of simulated log returns; bottom panels present the empirical densities of the extended jump test under Heston-A specification with $\Delta = 1sec$ and $M = 2500$. Grey area is the rejection region at the $\alpha = 1\%$ nominal significance level according to the Gumbel density.

The top panels of Figure 2.6 present the histograms of simulated returns, with a microstructure noise of $\varpi = 0.004$ and the following three levels of rounding. The left panel uses a light rounding of prices up to multiples of $g = 0.001$, the middle panel exhibits moderate rounding ($g = 0.005$) and the right panel displays severe rounding up to multiples of $g = 0.01$. The rounding takes place on the prices rather than on the log-prices. The support points of the log-returns vary therefore with the level of the price process.

The bottom panels of Figure 2.6 present the simulated distributions of the extended test statistic. We learn that with light rounding, the test statistic is consistent with the asymptotic Gumbel distribution as obtained in panel *i-b)*. Increasing the rounding to

$g = 0.005$ leads to a small size distortion, see panel *ii-b)*. However, when rounding effect is severe and price increments with mostly one-tick changes are observed, the simulated distribution of the test statistic in panel *iii-b)* is clearly affected. In this case, the use of a critical value based on the Gumbel density leads to an incorrectly sized test.

The presented simulation results depend on the combination of the average level of spot variance, sampling frequency, initial price level, and amount of microstructure noise. Each of those variables defines, together with the level of rounding g, the severity of discreteness in the return sequence. In our empirical study in Section 2.6, the rounding effect is light to moderate. We therefore expect a minor impact of rounding on our test statistic.

Effective power and size

We investigate effective power and size of the extended LMJ statistic (2.26) based on $\widetilde{\mathcal{L}}(t_i)$ in (2.29). The effective power shows how many jumps are correctly identified, that is 1 minus the frequency of failure to detect an actual jump. On the contrary, the effective size is interpreted as the probability at each t_i of a spurious jump detection that asymptotically equals $\alpha\% \times \Delta$. The more volatile the volatility process is, the more difficult jump detection becomes, see Lee and Mykland (2008). Our focus here centers on how sparse sampling, microstructure noise and leverage affect detection of jumps. For brevity, we confine ourselves to the Heston-A specification.

Table 2.2 presents simulation results for the effective power of applying the test to each individual price increment. As price observations are sampled more sparsely and Δ increases, the power to detect smaller-sized jumps decreases rapidly. For instance, for $\varpi = 0.001$ and $\rho = 0$, the extended test detects 74.38% of all jumps with size $\bar{\kappa} = 0.05$ at $\Delta = 1sec$, while only 0.06% at $\Delta = 5min$. The jumps from $\bar{\kappa} = 0.5$ are well detected at $\Delta = 5min$. This is not surprising given that such jumps contribute around 11% to total price variation. Additional microstructure noise has an adverse effect on detecting small jumps. When ϖ increases from 0.001 to 0.004, only 0.02% of all jumps are detected instead of the original 74.38%. In case $\varpi = 0.004$, the smallest jumps are not distinguishable from the innovations coming from the continuous part of the process. For the smaller jumps from $\bar{\kappa} = 0.05$ and $\bar{\kappa} = 0.15$, and large noise power increases initially with increase of Δ as microstructure noise is mitigated. Later it decreases again due to sparse sampling.

Table 2.2 also reports the daily effective size, defined as frequency of spuriously detected jumps divided by Δ. As long as the data is sampled at fine intervals, the increase of microstructure noise does not affect the effective size of extended test statistic. When we lower the sampling frequency, the estimate of the stochastic component in spot volatility becomes too smooth to capture high-frequency volatility movements. This leads to both power and size distortions.

We do not find a noticeable impact of leverage on the effective power and size of the

Table 2.2: *Power and size of extended test statistics*

	$\Delta = 1sec$	$\Delta = 5sec$	$\Delta = 30sec$	$\Delta = 1min$	$\Delta = 5min$
			$\varpi = 0.001$		
$\rho = 0$					
$\bar{\kappa} = 0.05$	0.7438 [0.009]	0.3651 [0.009]	0.0209 [0.014]	0.0044 [0.012]	0.0006 [0.038]
$\bar{\kappa} = 0.15$	1.0000 [0.006]	1.0000 [0.009]	0.9761 [0.013]	0.7273 [0.013]	0.0761 [0.028]
$\bar{\kappa} = 0.25$	1.0000 [0.007]	1.0000 [0.009]	0.9998 [0.009]	0.9993 [0.014]	0.4857 [0.027]
$\bar{\kappa} = 0.5$	1.0000 [0.010]	1.0000 [0.007]	0.9998 [0.014]	0.9998 [0.015]	0.9951 [0.020]
$\rho = -0.75$					
$\bar{\kappa} = 0.05$	0.7463 [0.009]	0.3647 [0.009]	0.0232 [0.016]	0.0053 [0.014]	0.0004 [0.038]
$\bar{\kappa} = 0.15$	1.0000 [0.006]	1.0000 [0.009]	0.9759 [0.014]	0.7263 [0.014]	0.0774 [0.028]
$\bar{\kappa} = 0.25$	1.0000 [0.006]	1.0000 [0.008]	0.9998 [0.009]	0.9993 [0.015]	0.4951 [0.025]
$\bar{\kappa} = 0.5$	1.0000 [0.012]	1.0000 [0.006]	0.9998 [0.015]	0.9998 [0.014]	0.9960 [0.019]
			$\varpi = 0.004$		
$\rho = 0$					
$\bar{\kappa} = 0.05$	0.0002 [0.0104]	0.0001 [0.006]	0.0014 [0.006]	0.0004 [0.007]	0.0006 [0.064]
$\bar{\kappa} = 0.15$	0.4190 [0.005]	0.4569 [0.006]	0.3179 [0.006]	0.1890 [0.009]	0.0581 [0.052]
$\bar{\kappa} = 0.25$	0.9779 [0.008]	0.9821 [0.006]	0.9636 [0.005]	0.9156 [0.005]	0.4020 [0.047]
$\bar{\kappa} = 0.5$	1.0000 [0.009]	1.0000 [0.006]	0.9998 [0.006]	0.9998 [0.004]	0.9907 [0.045]
$\rho = -0.75$					
$\bar{\kappa} = 0.05$	0.0002 [0.0104]	0.0001 [0.006]	0.0014 [0.006]	0.0003 [0.009]	0.0005 [0.069]
$\bar{\kappa} = 0.15$	0.4074 [0.005]	0.4604 [0.006]	0.3198 [0.005]	0.1896 [0.008]	0.0522 [0.060]
$\bar{\kappa} = 0.25$	0.9705 [0.007]	0.9843 [0.007]	0.9666 [0.004]	0.9130 [0.005]	0.4092 [0.054]
$\bar{\kappa} = 0.5$	1.0000 [0.009]	1.0000 [0.006]	0.9998 [0.006]	0.9998 [0.004]	0.9924 [0.054]

Note: Heston-A specification. Significance level $\alpha = 1\%$, thus $\alpha^* = 4.6001$. Effective power equals 1 minus frequency of failure to detect an actual simulated jump. Effective daily size between square brackets equals probability of detecting a spurious jump times number of observations per day at the sampling frequency. We set $h_n = n^{0.8}$, $\phi = 0.95$ and $M = 2500$.

extended test statistic based on $\widetilde{\mathcal{L}}(t_i)$ for data sampled at high frequencies. When Δ increases, the size is however more distorted. For instance, for $\Delta = 30sec$ and $\rho = 0$, the empirical daily size of a test equals 0.014 instead of the theoretical $\alpha = 0.01$, while it increases to 0.016 when $\rho = -0.75$. In terms of the effective size, it means an increase of the probability of a spurious jump detection at each tested time t_i from 4.86 $\times 10^{-6}$ to 5.56 $\times 10^{-6}$, instead of the theoretical $\alpha \times \Delta$ of 3.47 $\times 10^{-6}$.

2.6 Testing for Jumps Empirically

2.6.1 EUR/USD exchange rate data

In our empirical study we analyze the EUR/USD exchange rate data from July 2 to December 31, 2007.[2] The data time span is six months with $D = 131$ working days. The transaction price data at the sampling frequencies of 1, 5, and 15 minutes is provided by Disk Trading[3] and measures the price of 1 Euro expressed in U.S. Dollars with a precision of $g = 0.0001$. The time set is GMT+1. The percentage returns of the three price series are computed by taking their first differences in logs and multiplying these by 100.

Table 2.3: *Descriptive statistics*

Δ	mean	std.dev.	skewness	kurtosis	ACF_1	#obs
$1min$	3.60e-05	0.014	-0.083	16.954	-0.176 [±0.0046]	187816
$5min$	1.84e-04	0.027	0.212	22.045	-0.081 [±0.0103]	37670
$15min$	5.44e-04	0.044	0.234	14.366	-0.043 [±0.0178]	12571

Note: Descriptive statistics of EUR/USD returns. ACF_1 is the first order autocorrelation of raw returns; associated critical range is between square parentheses.

Table 2.3 presents the descriptive statistics of computed returns. The excess kurtosis of the return series indicates the presence of time-varying volatility and/or jumps. The first order autocorrelation of raw returns is negative and significant at all frequencies. This finding is a typical feature of price series contaminated with microstructure noise.

2.6.2 Estimation of diurnal volatility patterns

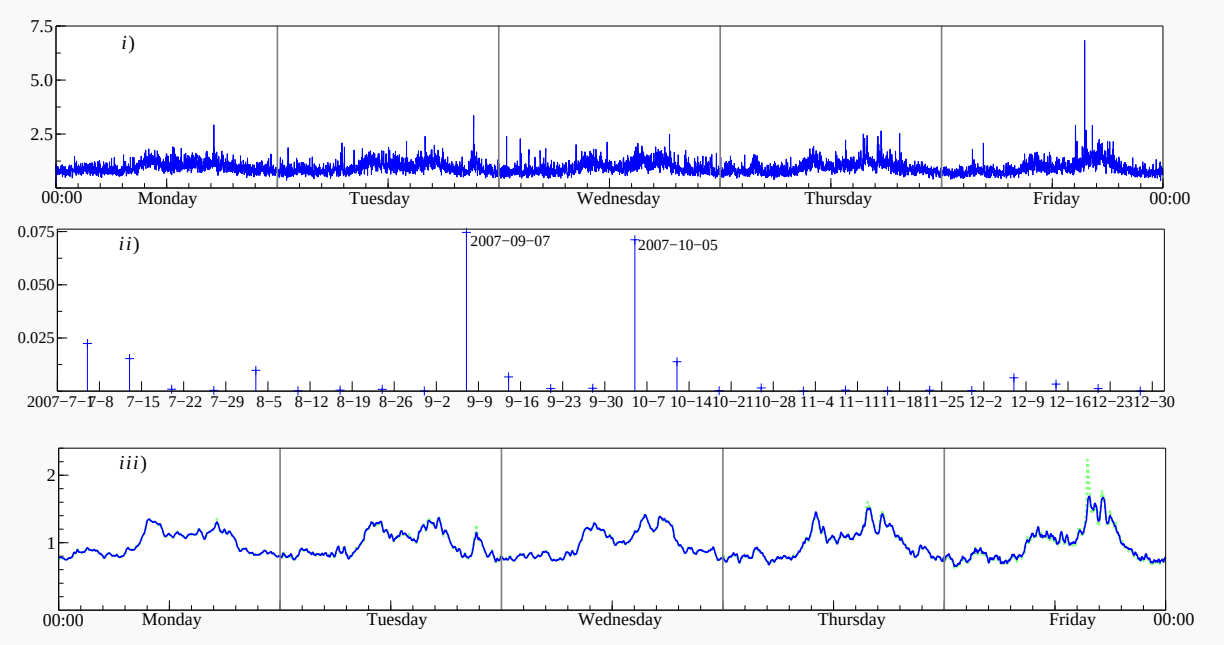
Figure 2.7 presents diurnal shapes estimated with the benchmark method (Section 2.3.3) and our robust procedure (Section 2.3.4). Panel *i)* presents the benchmark $\widehat{p}_{t_i}$ from (2.21) at $\Delta = 1min$ sampling frequency. Diurnal shapes are relatively similar for different days of the week. The main differences arise in the time period of 14.00-17.00, when markets in the US and Europe overlap. The numerous peaks in the return series correspond to various scheduled news announcements. European interest rate reports and macroeconomic releases appear on Tuesday and Wednesday whereas US macroeconomic news and interest rates decisions are released on Tuesday, Thursday and Friday.

There are numerous outliers in the benchmark $\widehat{p}_{t_i}$. To study whether these are caused by single distortions or are a part of the diurnal shape, panel *ii)* of Figure 2.7 presents the

[2]Our empirical study was also conducted on GBP/USD, JPY/USD and CHF/USD exchange rates. The results are relatively similar, thus for brevity only results for most heavily traded EUR/USD rate are presented.

[3]Source: `http://disktrading.is99.com`.

Figure 2.7: *Intraday periodic patterns*



Note: Periodicity extraction for the EUR/USD returns sampled at $\Delta = 1min$ frequency. Panel *i)* presents square root of cross-sectional variance (2.21); *ii)* presents absolute price increments of consecutive Fridays in the sample in time interval 15:00-15.31; *iii)* presents estimates of square root of a periodic factor (2.23) based on OWM-N (dotted line) and OWM-t (solid line).

absolute price increments in the time interval 15.30-15.31 of all Fridays. Two large return values appear on September, 7 and October, 5, which are the first Fridays of these months. Both days correspond to news surprises on payroll data from the US labor market survey.[4] At this particular intraday time, $\widehat{p}_{t_i}$ is affected by those two large price increments. To be robust against such news surprises, the estimation method of Section 2.3.4 is used. Panel *iii)* presents the estimated diurnal patterns based on weighting as in (2.23) with the use of observation disturbances from Gaussian (OWM-N) and Student's t (OWN-t) densities. The resulting patterns are much smoother and clearly display repetitive low and high intraday volatility periods.

Table 2.4 presents the likelihood-based estimation results of the parameters in the periodicity model (2.24)-(2.25). The largest outliers in this sample are on Thursday and

[4]The non-farm payrolls, released by the Bureau of Labor Statistics of the US Department of Labor, measure the number of people on the payrolls of all non-agricultural businesses. At the FX market a high reading is seen as positive for the US dollar, while a low reading is seen as negative. On September 7, the actual value of nonfarm payrolls (for August) was 93k, way below expected by market averaged at 108k. On October 5, the actual value of nonfarm payrolls (for September) was 96k, below expected by market averaged at 108k. On October 5, the actual value of nonfarm payrolls (for September) was 96k, below expected by market averaged at 100k. According to the economic calendar data at `http://www.fxstreet.com/` these came as market surprises.

Friday, which is reflected by low values for the degree-of-freedom parameter in the Student's t distribution ($\nu = 6.52$ and $\nu = 4.20$ respectively). Also, the excess kurtosis, $\frac{6}{\nu-4}$, is remarkably high for these days, especially on Friday (≈ 29.4). Given the low value of the estimates of ν and numerous outliers in the benchmark $\widehat{p}_{t_i}$, the robust method with Student's t observation disturbances is preferred.

Table 2.4: *Maximum likelihood estimates of the observation weight method*

	OWM-N		OWM-t		
	σ^2_ζ	σ^2_η	σ^2_ζ	σ^2_η	ν
Monday	0.0376	0.0003	0.0377	0.0003	14.1876
	(0.0015)	(0.0001)	(0.0017)	(0.0001)	(5.7772)
Tuesday	0.0398	0.0007	0.0404	0.0005	7.4998
	(0.0017)	(0.0002)	(0.0022)	(0.0002)	(1.4718)
Wednesday	0.0403	0.0004	0.0403	0.0004	12.2512
	(0.0016)	(0.0001)	(0.0018)	(0.0001)	(3.6748)
Thursday	0.0378	0.0007	0.0377	0.0006	6.5195
	(0.0015)	(0.0002)	(0.0021)	(0.0001)	(1.0686)
Friday	0.0524	0.0021	0.0437	0.0008	4.2041
	(0.0023)	(0.0005)	(0.0036)	(0.0002)	(0.5175)

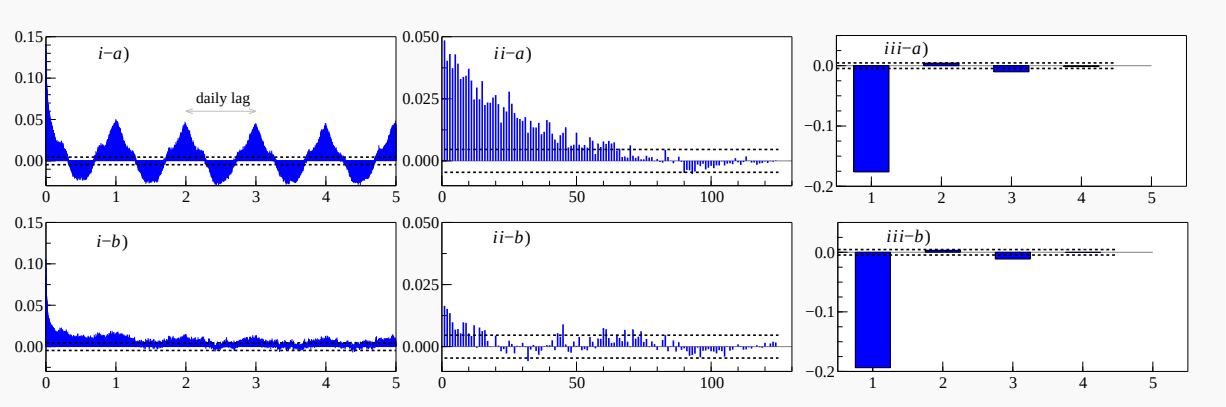
Note: Estimates of parameters of the periodicity model (2.24)-(2.25) for the EUR/USD rate at $\Delta = 1min$ sampling frequency. Standard errors between the brackets.

Panels *i-a)-iii-a)* in Figure 2.8 present the ACFs for absolute and raw returns at $\Delta = 1min$. The U-shape pattern recurring at a daily frequency in the autocorrelations of absolute returns can be seen in panel *i-a)*. Due to periodicity, the ACF of absolute returns is very persistent from day to day as shown in panel *ii-a)*. Panels *i-b)-iii-b)* present the ACFs after the diurnal pattern (robustly estimated) is removed. Periodicity is not perfectly extracted, but the ACF of absolute returns shown in panel *i-b)* has no clear recurring U-shaped cycles as comparing panels *i-b)* and *i-a)*. Moreover, the ACF is not that persistent, comparing panel *ii-b)* to panel *ii-a)*. The significant negative first order serial correlation seen in panels *iii-a)* and *iii-b)* is due to microstructure noise, and is not affected by removing the diurnal shape.

2.6.3 Spot measures

Figure 2.9 presents the returns (after robust adjustment for periodicity) and the estimated path of the stochastic component in the spot volatility based on the pre-averaged estimator (2.17) at $\Delta = 1min$. Estimates of σ^2_t are based on averaging over a set of local windows with $h_n = n^\beta$ with $\beta \in [0.52; 0.54, \ldots, 1]$. In panel *i)* of Figure 2.9, several return clusters (e.g., in August or November) can be recognized. These clusters correspond to the increased level of volatility $\widehat{\sigma}_{d+t_i}$; see panel *ii)*. Overall, the level of spot volatility increases and is possibly a sign of the start of the financial crisis in 2007. Estimated spot volatility reaches its maximum at the beginning of November 2007. The microstructure noise variance estimator of (2.11) is based on returns over one day. Panel *iii)* of Figure 2.9 displays 24-hour rolling window estimates of $\widehat{\varpi}_{d+t_i}$. It is found that the microstructure noise standard

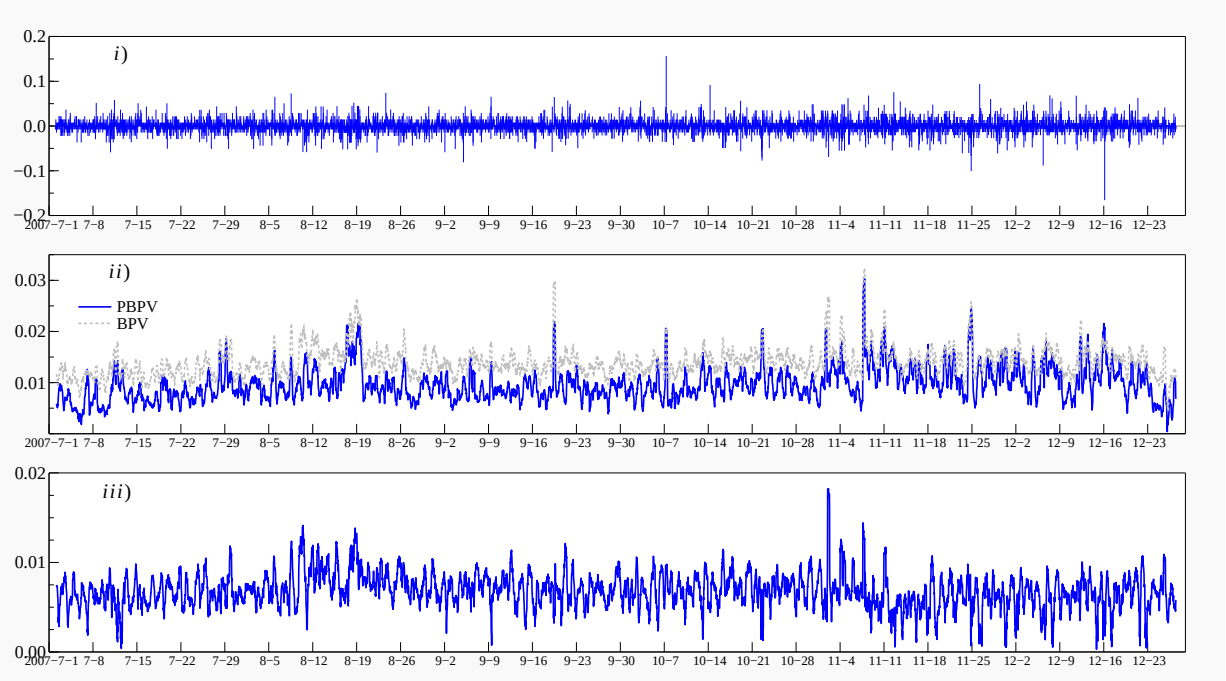
Figure 2.8: *Properties of returns*



Note: Autocorrelation functions of EUR/USD returns. Panels *i)* shows ACF of absolute returns displayed for 5 days; *ii)* ACF of absolute returns as in panel *i)* now showing only values at daily lags; *iii)* ACF of raw returns for 4 lags. Panels *a)* raw returns; *b)* returns adjusted for variance diurnal shape using OWM-*t*.

deviation is relatively high at the beginning of the sample, while lower towards the end of the sample. As a result, the estimated noise-to-signal ratio in (2.18) also varies considerably over time. Its value (not displayed) starts from around 80% at the beginning of the sample period, decreasing to 40-50% at the end of the sample.

Figure 2.9: *Spot measures*

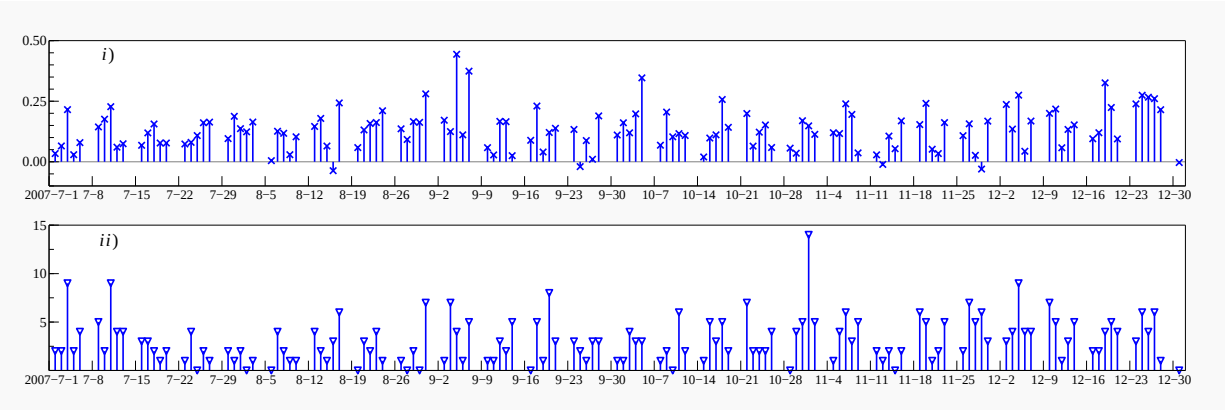


Note: Panel *i)* shows EUR/USD periodically adjusted returns at $\Delta = 1min$ frequency; *ii)* estimated stochastic component $\widehat{\sigma}_{d+t_i}$ (times $\sqrt{\Delta}$) of the spot volatility path; *iii)* estimated time-varying volatility of microstructure noise $\widehat{\varpi}_{d+t_i}$.

2.6.4 Aggregate evidence of jumps

We first report daily estimates $\widehat{\text{PBPV}}_d(2,0)$ and $\widehat{\text{PBPV}}_d(1,1)$ for $d = 1, \ldots, 131$. The relative jump variation estimate $\widehat{\text{RJV}}_d$ indicates the presence of a jump component in the price process. Given the sampling frequency $\Delta = 1min$, the average of $\widehat{\text{RJV}}_d$ series equals 13% (16% and 22% at $\Delta = 5min$ and $\Delta = 15min$, respectively). Mean estimates of RJV increase with an increasing Δ, but the accuracies of IV and QV estimates decrease. Using realized measures that are not robust to microstructure noise, Huang and Tauchen (2005) find that jumps contribute in around 7% in the S&P index at $\Delta = 5min$. Barndorff-Nielsen and Shephard (2006) find that jumps contribute between 5% and 22% to total price variation dependently on the exchange rate at frequencies ranging from $\Delta = 5min$ to $\Delta = 1h$.

Figure 2.10: *Relative jump contribution and number of identified jumps*



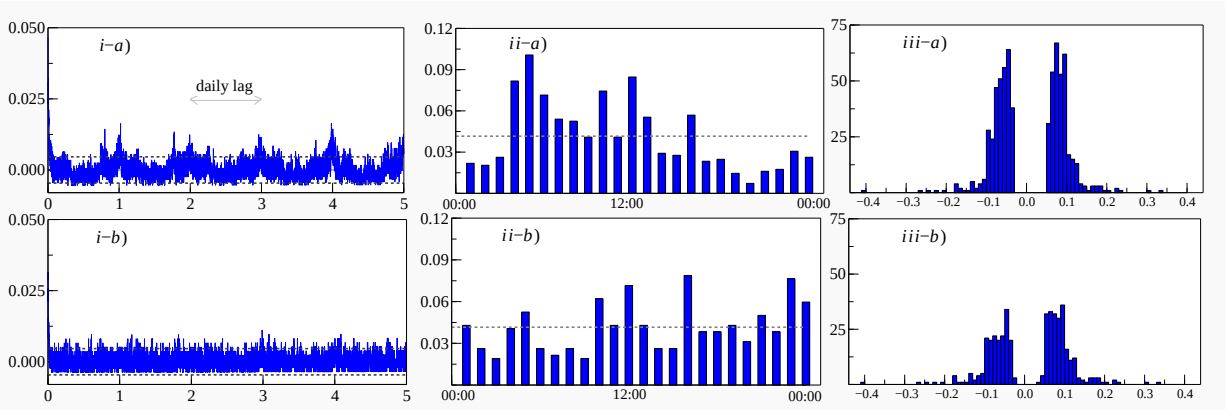
Note: Relative jump variation (2.13) in panel *i*) and the number of significant jumps in panel *ii*). Significant jumps based on the LMJ test (2.26) with $\widetilde{\mathcal{L}}(t_i)$ in (2.29) at the $\alpha = 1\%$ significance level.

Figure 2.10 presents the sequence of RJV estimates and the number of significant jumps over the sample period at $\Delta = 1min$. There are numerous days when $\widehat{\text{RJV}}_d$ exceeds 25-30% as well as when it is roughly zero. Using the extended LMJ statistic we find more jumps at the higher frequency.[5] At $\Delta = 1min$, in total 419 jumps are detected at $\alpha = 1\%$, so that 3 jumps occur on average per trading day. When $\Delta = 5min$ 132 jumps are detected, which means that 1 jump occurs on average per day. Finally, at $\Delta = 15min$ sampling frequency only 45 jumps are detected, which results in 1 jump on average every 3 days. We find a maximum of 2 jumps (July 18, September 5, October 5 and December 27) at $\Delta = 15min$, while we find a maximum of 14 jumps (November 1) at $\Delta = 1min$. The significant jumps at $\Delta = 1min$ have an average absolute size of 8.4 basis points (bps), while at $\Delta = 5min$

[5]Note that the LMJ statistic was compared to the critical value resulting from the Gaussian assumption. A separate simulation exercise mimicking the amount of rounding occurring at the present data set indicated that this critical value is slightly conservative.

($\Delta = 15min$) the average size is of 15.6bps (24.4bps).

Figure 2.11: *Characteristics of identified jumps*



Note: Autocorrelation function of jump indicators in panel *i)* for 5 days; timing aggregated to 1 hour in panel *ii)*; histogram of timing of significant jumps in panel *iii)*, at $\Delta = 1min$. Panels *a)* for original LMJ test based on $\mathcal{L}(t_i)$ in (2.28); *b)* for the extended LMJ test based on $\widetilde{\mathcal{L}}(t_i)$ in (2.29).

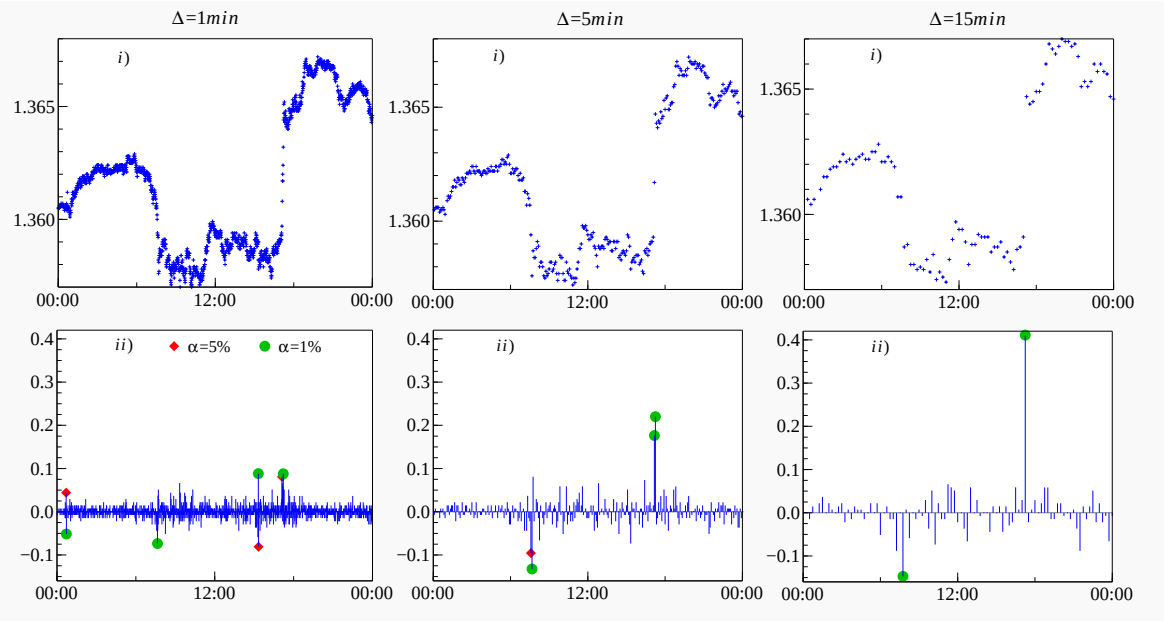
Figure 2.11 presents the dependence structure, timing and histogram of the detected jumps. We compare the LMJ test statistic based on (2.28) and on the extended statistic (2.29). Since the original LMJ test does not account for diurnal patterns, the autocorrelation structure indicates the periodical dependence of the detected jumps; see the autocorrelation clusters at daily time intervals in panel *i-a)* of Figure 2.11. The LMJ statistic detects more (less) jumps when deterministic volatility is higher (lower); see panel *ii-a)*. It causes spurious time dependence of jumps. Jumps detected with the extended statistic are not periodically dependent and are more equally distributed over intraday time as seen in *i-b)* and *ii-b)*, respectively, of Figure 2.11. Over this sample period, there were more positive jumps than negative as observed in panel *iii-b)* of Figure 2.11. Given that the US dollar was depreciating in those months, this result is not surprising. The extended test statistic detects more smaller-sized jumps than the original test; compare the histogram in panel *iii-b)* to *iii-a)*. The original test detects more jumps but in the Monte Carlo study it is shown that the presence of diurnal patterns leads to an overrejection of the null hypothesis of no jumps. Too many jumps are spuriously detected due to the periodic increase of return volatility that is not captured when we estimate the two spot variance components jointly.

2.6.5 Jumps at specific days

The LMJ jump test statistic yields the timing, size and direction of the detected jumps. We study in some detail the outcome of applying the extended test for two specific days. On the one hand, we select Wednesday, September 5, when $\widehat{\mathrm{RJV}}_{2007/09/05}$ reaches its maximum

of 44.4% at $\Delta = 1min$, see Figure 2.10. On the other hand, we choose Thursday, November 1, when a maximum number of 14 jumps is found at $\Delta = 1min$, while $\widehat{\text{RJV}}_{2007/11/01}$ equals 14.8% at this frequency. Figures 2.12a and 2.12b zoom in on both days.

Figure 2.12a: *Case study for September 5, 2007.*

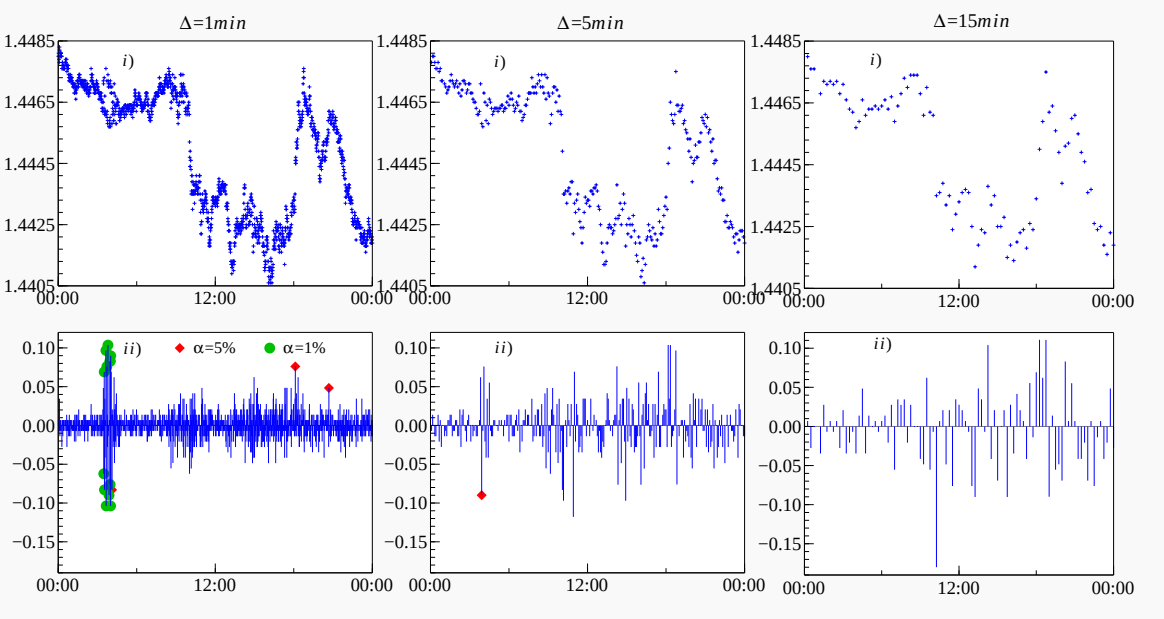


Note: Panels *i)* present the EUR/USD exchange rate in level; *ii)* EUR/USD returns and significant jumps based on the extended LMJ test with $\widetilde{\mathcal{L}}(t_i)$ in (2.29) (at $\alpha = 1\%$ marked with circles; at $\alpha = 5\%$ marked with diamonds).

Selected days in this case study correspond to scheduled macroeconomic announcements relevant for the EUR/USD exchange rate, these are listed in Table 2.5. On September 5 there are 4 jumps detected at $\Delta = 1min$ frequency at 00:44, 07:38, 15:18 and 17:12. At the 5% significance level, three additional jumps are found at 00:43, 15:19 and 17:10; see panel *ii)* of Figure 2.12a. These time intervals correspond to high intraday volatility time periods, when markets open or overlap. The morning European news items were positive for the EUR/USD rate on that day. It is seen from Figure 2.12a that the EUR/USD rate started to appreciate dramatically in the afternoon when the US readings, presented in Table 2.5, were much worse than expected. The sizes (in absolute terms) of the detected jumps were from 7 to 10bps (or from 6bps at $\alpha = 5\%$), which corresponds to around 0.07-0.1% change of the fundamental within a $1min$ interval. When lowering the sampling frequency to $\Delta = 15min$ only 2 jumps are found to be significant, at 15:18 (-15bps) and at 17:12 (42bps) contributing around 51.6% to the total price variation as measured by the estimate of RJV.

On November 1 there are 14 (17) jumps found at $\Delta = 1min$ frequency at the 1% (5%) significance level. Interestingly, 12 of these jumps cluster in the time period between 03:28-04:06 when only Asian markets operate as seen in panel *ii)* of Figure 2.12b. These jumps

Figure 2.12b: *Case study for November 1, 2007.*



Note: See Figure 2.12a.

have an absolute size of around 10bps. We have no detailed information what happened in this time interval, but on this particular Thursday many US news announcements were to be released, see Table 2.5, which possibly caused some additional arrival of orders. At $\Delta = 1min$, $\widehat{\text{RJV}}_{2007/11/01}$ equals 14.8%, which indicates the presence of a jump component. When lowering the sampling frequency to $\Delta = 15min$, we find no significant jumps even at the $\alpha = 5\%$ significance level, although the estimate of RJV equals 17.7% at this frequency and the largest absolute price increment is of 18bps. It indicates that even when estimates of QV and IV differ considerably (and hence indicate the occurrence of jumps) it may not be possible to point out which price increments are caused by jumps when the sampling frequency is low.

2.7 Summary and Conclusions

In this chapter we introduce a new approach to the estimation of spot variance from high-frequency data. We decompose spot variance into stochastic and deterministic components. Estimation of the stochastic variance component is based on adopting a pre-averaged bipower variation measure. The pre-averaging approach allows us to estimate the stochastic component in the jump-diffusion framework with microstructure noise. As a result, we can apply the estimator without the need for mitigating noise effects through sparse sampling. Estimation of the deterministic variance component is based on the newly

Table 2.5: *News releases.*

September 5, 2007						
Time	Area	Event	Actual	Consensus	Previous	+/-
08:55	DE	PMI Services (Aug)	59.8	58.1	58.5	+
09:00	EMU	PMI Services (Aug)	58.0	57.9	58.3	
10:00	EMU	Retail Sales (MoM) (Jul)	0.4%	0.3%	0.6%	+
10:00	EMU	Retail Sales (YoY) (Jul)	1.3%	1.1%	1.0%	+
13:15	US	ADP Employment Change (Aug)	27k	80k	48k	+
15:00	US	Pending Home Sales (MoM) (Jul)	-12.2%	-2.0%	5.0%	+
19:00	US	Fed's Beige Book				

November 1, 2007						
Time	Area	Event	Actual	Consensus	Previous	+/-
13:30	US	Core PCE - Prices Index (MoM) (Sep)	0.2%	0.2%	0.1%	
13:30	US	Core PCE - Prices Index (YoY) (Sep)	1.9%	1.8%	1.8%	-
13:30	US	Initial Jobless Claims (Oct 27)	330k	330k	331k	
13:30	US	Personal Income (MoM) (Sep)	0.4%	0.4%	0.4%	
13:30	US	Personal Spending (Sep)	0.3	0.4	0.6	
15:00	US	ISM Manufacturing (Oct)	50.9	51.4	52.0	+
15:00	US	ISM Prices Paid (Oct)	63	64	59	

Note: Important news releases on September 5, and November 1, 2007. The + (-) denotes a positive (negative) reading for the EUR/USD exchange rate. DE - Germany, EMU - European Monetary Union, PMI - Purchasing Manager Index, PCE - Personal Consumption Expenditure.
Source: `http://www.fxstreet.com/`.

developed semi-parametric approach. We smooth the cross-sectional variances using a robust weighting scheme. In a finite sample Monte Carlo study we show that the underlying spot variance path can be estimated accurately as long as the price observations are sampled at a sufficiently high-frequency, indeed even in practical situations where microstructure noise, rounding effects and jumps are present or when leverage effects confound the variance process.

Particular attention is given to the testing for jumps since our framework with microstructure noise offers an extension to the jump test statistic introduced by Lee and Mykland (2008) and modified later by Boudt et al. (2011). The Monte Carlo study has shown that the distribution of our extended test statistic is not distorted by periodic patterns, leverage effects or microstructure noise. Also, our extended jump test improves on the detection of smaller-sized jumps, which, however, also depends on the magnitude of microstructure noise contamination. With large noise, the smallest jumps cannot be distinguished from microstructure frictions, which is a theoretical result and not due to a failure of the testing method. The power of our extended test decreases when the sampling is less frequent but we should emphasize that our framework alleviates the need for sparse sampling. The asymptotic density of our extended test statistic relies on the assumption that the microstructure noise is Gaussian. Since at the highest sampling frequency the prices live on a discrete grid, this assumption is obviously too restrictive. However, at a moderate sampling frequency and in particular for our dataset with a grid-size of 0.0001 we found that such discreteness has little or no impact on the distribution of the test statistic.

The empirical illustration focuses on spot volatility and jumps in the EUR/USD exchange rate series for a period of 2007 when the financial crisis became apparent. In this period the spot volatility has increased slowly while the microstructure noise has become less important in the later months. We provide clear evidence of discontinuities in the exchange rate series. More jumps are detected during days with important macro news announcements and from a dataset that is sampled at a higher frequency. When the periodic volatility factor is extracted and removed, we do not find periodic time dependence in the realized jump series and the number of jumps is also reduced. We therefore conclude that our modifications for estimating spot volatility and jump testing are effective and can be successfully applied in studies of financial markets based on intraday high-frequency data.

Chapter 3

Quantile-based realized measure of variation: A new test for outlying observations in financial data

Abstract: In this chapter we consider the moment- and quantile-based estimators of power variation as measured by σ^r, with σ being the scale parameter and r a positive integer. We derive the joint asymptotic distribution under the normal distribution. With the limiting distribution, we introduce a new class of test statistics designed to detect abnormal observations. The power of the tests generally increases with r, as higher moments accentuate outlying observations. In addition, we can statistically judge the impact of outlying observations on the rth power variation of interest. We apply the tests to detect jumps in financial data as implied by the Merton (1980) model. Simulation results show that the finite sample properties of the test depend on the order r and the quantiles applied. Our novel test for jumps is found to be more powerful than widely used alternatives for sufficiently large sample sizes. An empirical illustration for high-frequency equity data suggests that jumps can be more prevalent than inferred from the second or third moment of the data.

3.1 Introduction

It is widely acknowledged that financial returns are non-normal as their distribution typically exhibits some degree of asymmetry and excess kurtosis. The non-Gaussian features of the return distribution can be explained by stochastic volatility and/or the presence of jumps, which are defined as abnormal price increments; see Das and Sundaram (1999). There is considerable empirical evidence that volatility of assets is time-varying but also

that jumps are present in the underlying price process, see e.g. Carr and Wu (2003), Barndorff-Nielsen and Shephard (2006), Andersen, Bollerslev, and Dobrev (2007), Lee and Mykland (2008), Jiang and Oomen (2008), Aït-Sahalia and Jacod (2009) and Bos, Janus, and Koopman (2009). For any particular series, there are however still two important issues: First one has to decide whether abnormal returns are indeed present in this series, and secondly the impact, if any, of these jumps on the variation of the price process must be assessed.

Most of the existing tests designed to detect jumps decompose quadratic variation of the price process into a continuous part, the so-called integrated variance, and a part due to jumps. Such tests identify, through the second moment, the presence of jumps. They however do not quantify the impact of jumps on other measurements beyond integrated variance. The primary goal of this chapter is to provide a statistical method making use of arbitrary powers of return data, leading to formal inference on the impact of potential abnormal observations also on other moments of the data. For instance, even though some jumps might not be detectable by tests built on squared returns, their possible presence might still affect higher moments.

The primary motivation for our study is the increasing use of higher moments of the data for a broad scope of financial applications; for a comprehensive review we refer to Jondeau, Poon, and Rockinger (2007). When going beyond variance as a measure of risk, the standard measures of skewness and kurtosis are mostly applied. Likewise, the integrated quarticity, a variance of integrated variance, can be applied as additional measure of risk and hence it is of considerable interest in financial applications, see Corsi, Mittnik, Pigorsch, and Pigorsch (2008). Since it is widely agreed that unconditional non-Gaussianity of financial returns is either due to time varying volatility and/or jumps, the standard moment-based measures quantify a total risk. However, it is of natural interest in financial economics to distinguish between the volatility and jump risk, as their implications differ from portfolio and risk management to option or bond pricing and hedging, see e.g. Bakshi, Cao, and Chen (2000) and Johannes (2004). To separate the jump risk, we consider two estimators of the power variation jointly, where one estimator is robust to jumps and the other is not.

If the underlying return data is conditionally normally distributed (or following any other distribution known up to location and scale), the qth sample quantile equals the mean plus the standard deviation times the inverse of the standard normal cumulative distribution function at probability q. It is possible to invert this relationship and solve for the standard deviation, but also for mean. Such quantile-based measures for volatility for normal data date back to Pearson (1920). The appealing feature of the quantile-based approach is the robustness to extreme observations, provided that the selected quantiles are sufficiently far away from the extreme tails. Christensen, Oomen, and Podolskij (2010) make use of this feature and propose to use sample quantiles to obtain a robust-to-jumps

measure of integrated variance. Relative to the moment-based measure of standard deviation, efficiency of the quantile-based measure depends on the choice of sample quantiles employed. Mosteller (1946) was first to suggest combining more quantiles to improve efficiency of the quantile-based measure of standard deviation, such that it can approach the efficiency bound of the maximum likelihood estimator, while still being robust to extreme observations. Having the moment- and quantile-based measures of standard deviation, it is natural to investigate their joint distribution. The answer is given recently by DasGupta and Haff (2006) who provided the joint asymptotic distribution of the inter-quartile range and standard deviation, and proposed a diagnostic for testing whether the underlying data is normal. Their approach however is primarily based on standard deviation (the first order power variation) and inter-quantile range, while in this chapter we exploit higher orders of power variation as measured by moment- and quantile-based measures. This will prove to provide additional power for detecting abnormal observations.

Our methodological contribution can be summarized as follows. We introduce the quantile-based measure of the rth order power variation of univariate data as measured by σ^r, with σ being the scale parameter and r being a positive integer, under the assumption of normality. We provide optimal quantiles for the estimator under the minimum asymptotic variance criterium, and we provide weights to optimally combine multiple quantile-based sub-estimators with the aim to improve efficiency further. For our purpose, we consider the moment-based estimator of σ^r and, as a central contribution of this chapter, we derive the joint limiting distribution of the two estimators assuming normality of the underlying return distribution. Contrary to the quantile-based estimator, the moment-based estimator is extremely sensitive to outliers. This proves useful in construction of a Hausman (1978) type test designed to detect outlying observations. The null hypothesis is that abnormal observations are not present in the data.

We use the quantile-based measure of variation for testing the occurrence of jumps, using an arbitrary order r of returns to accentuate possible outlying observations, measuring their effect on the power variation of interest. This contrasts our setup with the one of Christensen et al. (2010), who use a similar quantile-based variation measure over short time intervals, to combine these into a daily integrated variation measure.

In principle, our testing procedure can be applied to any constant variance dataset, where it is conjectured that the presence of a finite fraction of outliers leads to non-normality. In this chapter, we apply the test to detect jumps in financial return data and we compare its performance to widely applied alternatives: the test of Barndorff-Nielsen and Shephard (2006) and of Jiang and Oomen (2008). Our test is less general, as it requires a constant variance assumption. This test can therefore be considered when financial data is sampled over a suitably short time horizon (cf. Christensen et al., 2010) or on a suitably adapted time scale. In our application based on high-frequency data, we use a sampling scheme that for each trading day forms a series of intraday returns of equal variance. Our

simulation results show that our test has good finite sample properties and that it can be more powerful than the alternatives, depending on the sample size and the order r. An extensive empirical illustration for fifteen NYSE equities suggests that jumps can occur more frequent than would be inferred from only the second moment of the data.

The structure of the chapter is as follows. In Section 3.2, we review the theory of sample quantiles, we set out our notation and we provide the quantile-based estimator of power variation. In Section 3.3, we provide the joint asymptotic distribution of the quantile- and moment-based measures of σ^r under normality. We propose also a new class of test statistics designed to detect abnormal observations. In Section 3.4, we study the behavior of the quantile-based estimator and resulting test statistics in a simulation study. In Section 3.5, we apply our theory to equity data, while Section 3.6 concludes. The proofs of the main results in this chapter are given in Appendix 3.A.1.

3.2 The Quantile-based Measurement of Variation

In this section, we briefly review the theory of sample quantiles. We next introduce the quantile-based estimator of σ^r for positive integer-valued r. Derivation of the estimator with its asymptotic distribution is carried out under the assumption of the normal distribution.

3.2.1 Moment-based measurement of variation

First, we introduce the standard moment-based estimator of σ^r for the normal data. Let the observations $Y = \{Y_i\}_{i=1}^N$ be independently and identically normally distributed with mean μ and standard deviation σ. We use the central absolute moments of the data and define the moment-based estimator of rth order power variation as

$$MPV_N^r = \frac{1}{b(r)} \frac{1}{N} \sum_{i=1}^{N} |Y_i - \mu_N|^r \qquad \text{with} \quad \mu_N = \frac{1}{N} \sum_{i=1}^{N} Y_i, \tag{3.1}$$

where

$$b(r) = \frac{2^{r/2} \Gamma\left((r+1)/2\right)}{\sqrt{\pi}} \tag{3.2}$$

is the normalizing term. As $N \to 0$, it follows that

$$N^{1/2} \left(MPV_N^r - \sigma^r \right) \xrightarrow{d} \mathcal{N}\left(0,\, \sigma^{2r} \frac{M(r)}{b^2(r)}\right),$$

where

$$M(r) = M_z^{2r} - (M_z^r)^2 \qquad \text{with} \qquad M_z^n = \int_{-\infty}^{\infty} |z|^n \phi(z) \mathrm{d}z, \tag{3.3}$$

where z denotes an *iid* standard normal variable and $n \in \mathbb{N}$.

Note that we specifically consider an explicit moment-based estimator of σ^r for any positive integer-valued r, as higher orders of r accentuate outliers and hence inflate their impact on a direct estimate of σ^r. Alternatively, one could have considered the maximum likelihood (ML) estimator,

$$\hat{\sigma}^2_{\text{ML}} = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \mu_N)^2, \qquad \text{with} \quad \hat{\sigma}^r_{ML} = (\hat{\sigma}^2_{ML})^{r/2}. \tag{3.4}$$

As $N \to \infty$ we have that $N^{1/2}(\hat{\sigma}^r_{\text{ML}} - \sigma^r) \xrightarrow{d} \mathcal{N}(0, \frac{r^2}{2}\sigma^{2r})$. Note that $\hat{\sigma}^2_{\text{ML}} \equiv MPV^2_N$, implying that the second order moment-based estimator attains the efficiency bound of the ML estimator.

Figure 3.1: *Efficiency of MPV relative to ML*

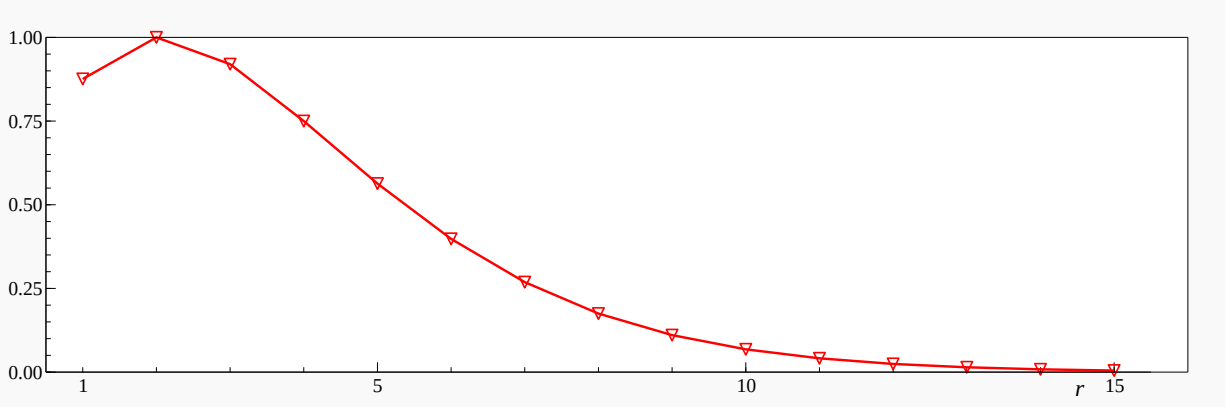


Figure 3.1 presents the efficiency of MPV^r_N relative to that of the corresponding ML estimator. While efficiency is still high for a small order of r, it declines dramatically once $r > 6$. As the MPV^r_N estimator is increasingly sensitive to outliers with increasing order of r, it still may prove valuable (and more so than the ML estimator) in constructing a test statistic aimed at detecting potential outlying observations.

3.2.2 Quantiles and order statistics

Let $X = \{X_i\}_{i=1}^N$ denote an independently and identically distributed (*iid*) sequence of random variables from a univariate continuous distribution with cumulative distribution function (*cdf*) F and probability density function (*pdf*) f. Denote the qth quantile by $Q(q) = F^{-1}(q)$ for $0 \leq q \leq 1$, implying that $q = \int_{-\infty}^{Q(q)} f(x)\mathrm{d}x$. For instance, for $q = 0.5$ we obtain the so-called median measure of the variable of interest.

The simplest approach of computing the quantile $Q(q)$ empirically is by using the order statistics directly,

$$Q_N^o(q) = X_{([qN])} \tag{3.5}$$

where $[\alpha]$ is the operator which rounds α to the nearest integer, and $X_{(k)}$ indicates the kth order statistic of X, such that $X_{(1)} \leq \ldots \leq X_{(k-1)} \leq X_{(k)} \leq X_{(k+1)} \leq \ldots \leq X_{(N)}$. If $q < \frac{1}{2N}$, then the order statistic $X_{(0)}$ would be needed; we define the 0th order statistic to be equal to $X_{(1)}$ for completing the above definition. Likewise, if (below) an order statistic with $k > N$ is requested, then this will be replaced by $X_{(N)}$.

Apart from this most basic definition of an empirical quantile, Hyndman and Fan (1996) defines a range of 9 different methods for calculating quantiles. For improved accuracy in cases of small samples, it is important to interpolate between order statistics, instead of rounding to the nearest one. Hence, in the remainder of this chapter we estimate the quantile as a weighted average between two adjacent order statistics,

$$Q_N(q) = wX_{(l)} + (1-w)X_{(l+1)}, \tag{3.6}$$

where the lower order statistic is chosen using $l = \lfloor (N+1)q \rfloor$ with $\lfloor \alpha \rfloor$ denoting the rounding down of α, and the weight is $w = l + 1 - (N+1)q$. With this setup (which corresponds to method 6 of Hyndman and Fan (1996)), the kth order statistic corresponds to quantile $k/(N+1)$, effectively spreading out the order statistics uniformly over the quantiles $1/(N+1), \ldots, N/(N+1)$.

It is well known that the empirical quantile (irrespective of the choice of definition (3.5) or (3.6)) is asymptotically normally distributed with

$$N^{1/2}\left(Q_N(q) - Q(q)\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{q(1-q)}{[f(Q(q))]^2}\right), \quad \text{as} \quad N \to \infty, \tag{3.7}$$

provided $f(Q(q)) = f(F^{-1}(q)) > 0$. Here $\xrightarrow{d}$ denotes convergence in distribution. Actually, rather than considering distribution of the single sample quantile, we might be interested in the joint asymptotic distribution of a finite set of sample quantiles, which is given in Theorem 3.2.1.

Theorem 3.2.1. *Define a p-dimensional vector* $\boldsymbol{q} = \{q_1, \ldots, q_p\}$ *with* $0 < q_1 < \ldots < q_p < 1$ *determining sample quantiles* $\mathcal{Q}_N(\boldsymbol{q}) = [Q_N(q_1), \ldots, Q_N(q_p)]'$ *and corresponding theoretical quantiles* $\mathcal{Q}(\boldsymbol{q}) = [Q(q_1), \ldots, Q(q_p)]'$. *The joint asymptotic distribution of multiple quantiles is given by the p-variate normal distribution,*

$$N^{1/2}\left(\mathcal{Q}_N(\boldsymbol{q}) - \mathcal{Q}(\boldsymbol{q})\right) \xrightarrow{d} \mathcal{N}\big(\mathbf{0}, \boldsymbol{\Sigma}(\boldsymbol{q})\big), \tag{3.8}$$

where the ijth element of the $p \times p$ covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{q})$ of quantiles is given by

$$\Sigma_{ij}(q_i, q_j) = \big(\boldsymbol{\Sigma}(\boldsymbol{q})\big)_{ij} = \big(\boldsymbol{\Sigma}(\boldsymbol{q})\big)_{ji} = \frac{q_i(1-q_j)}{f(Q(q_i))f(Q(q_j))}, \qquad \text{for} \quad i \leq j;\ i,j = 1,\ldots,p. \tag{3.9}$$

Proof. See Walker (1968).

For simplicity in notation (like below), the explicit dependence of $\boldsymbol{\Sigma}(\boldsymbol{q})$ on the *pdf* and inverse *cdf* is omitted.

Finally, in the next section we are interested in a function of sample quantiles and its associated limiting distribution. The following corollary will be used:

Corollary 3.2.2. *Let φ denote a function of real variables z that is fully differentiable. Let $\dot{\varphi}$ denote the first order derivative that is calculated at $z = \mathcal{Q}(\boldsymbol{q})$ and that is non-zero valued. By applying the delta-method to the joint asymptotic distribution of multiple quantiles $\mathcal{Q}_N(\boldsymbol{q})$ from Theorem 3.2.1, it follows that*

$$N^{1/2}\left(\varphi\big(\mathcal{Q}_N(\boldsymbol{q})\big) - \varphi\big(\mathcal{Q}(\boldsymbol{q})\big)\right) \xrightarrow{d} \mathcal{N}\big(0, \dot{\varphi}(\boldsymbol{q})'\,\boldsymbol{\Sigma}(\boldsymbol{q})\ \dot{\varphi}(\boldsymbol{q})\big). \tag{3.10}$$

The result of Corollary 3.2.2 has been used in many situations. For instance, Moors, Wagemakers, Coenen, Heuts, and Janssens (1996) apply it to derive the limiting distributions of quantile-based measures of skewness and kurtosis.

3.2.3 Quantile-based measurement of variation and its properties

Apart from the moment-based measures of data variation, many others have been developed. For an excellent history with a particular focus on early developments, see David (1998). For example, the inter-quantile range, i.e. the range between the quantiles of order q and $1-q$, is one of the oldest measures of dispersion of data. The inter-quantile range of any (fully specified, up to location and scale) symmetric distribution is a known multiple of the standard deviation. In the following, we will discuss results for the Gaussian distribution, as this is the distribution that will be of use in the empirical application in Section 3.5.

Assumption 1. *The observations $Y = \{Y_i\}_{i=1}^N$ are independently and identically normally distributed with mean μ and standard deviation σ.*

For such observations, there is a well-known relationship between the sample quantile and standard deviation given by

$$Q(q) = \mu + \sigma\Phi^{-1}(q), \tag{3.11}$$

where Φ^{-1} denotes the inverse *cdf* of the standard normal variable. If we consider another quantile, $(1-q)$th say due to symmetry, we can solve for the standard deviation even when the location μ is unknown, as in

$$\sigma = \frac{1}{c(q)}\Big[Q(q) - Q(1-q)\Big], \tag{3.12}$$

with $c(q) = \Phi^{-1}(q) - \Phi^{-1}(1-q) = 2\Phi^{-1}(q)$ the scaling term corresponding to the standard normal distribution. Using empirical quantiles, this delivers us our quantile-based first order power variation,

$$QPV_N^1 = \frac{1}{c(q)}\Big[Q_N(q) - Q_N(1-q)\Big]. \tag{3.13}$$

Such a quantile-based estimator of first order power variation for σ was first proposed by Pearson (1920). He proposed to use $q = 1-1/14 \approx 0.9286$, as a value of the quantile which would result in a relatively efficient estimator of σ. Benson (1949) examined the variance of (3.12) for different values of q and found that efficiency is greatest for $q = 0.9308$. Mosteller (1946) showed the asymptotic normality of the estimator (3.12) and he also suggested to combine multiple quantiles to improve the efficiency further. Ogawa (1951) further improved the quantile-based measure of σ by deriving a linear combination of a specified number of sub-estimators, based on symmetric quantiles as in (3.12), which are optimally weighted for maximum efficiency. All these authors stress the robustness to more extreme observations of the quantile-based approach which makes it appealing relative to the moment-based approach. Also, in practical situations, it might be often the case that normality is a good approximation to a central region of empirical density, say except for 1% or 3% of observations in each tail. In such case, the robustness property inherent to the quantile-based approach may prove extremely useful.

For our purpose, we more generally consider estimation of any positive integer-valued power of σ, i.e. σ^r for $r \in \{1, 2, \ldots\}$, using the sample quantiles of the data. Along with the moment-based estimator of σ^r, the quantile-based estimator will be used in Section 3.3 to construct a new class of test statistics for detecting outlying (or abnormal) observations. For that purpose, we introduce an estimator that is built on the rth power of the estimator in (3.12). In principle, we could consider estimation of σ^r for any $r > 0$, though this would not change the main results; hence we simplify and restrict ourselves to integer values of r.

Consider the general quantile-based estimator of rth order power variation,

$$QPV_N^r(q) = \frac{1}{c^r(q)} \sum_{j=0}^{r} (-1)^j \binom{r}{j} \big[Q_N(q)\big]^{r-j} \big[Q_N(1-q)\big]^j, \tag{3.14}$$

for $r \in \mathbb{N}_1 = \{1, 2, \ldots\}$ and $q \in (1/2,\, 1)$, where $\binom{r}{j}$ denotes the binomial coefficient. Since the quantile-based estimator (3.14) is constructed through a differentiable function of two

quantiles, the limiting distribution can be derived using the general results of Section 3.2.2. We obtain

$$N^{1/2}\left(QPV_N^r(q) - \sigma^r\right) \xrightarrow{d} \mathcal{N}\left(0,\, 2(1-q)(2q-1)\, d^2(q,r)\sigma^{2r}\right), \tag{3.15}$$

where

$$d(q,r) = \frac{\left[\Phi^{-1}(q)\right]^{r-1} \sum_{j=0}^{r} j\binom{r}{j}}{c^r(q)\phi(\Phi^{-1}(q))} = \frac{\sum_{j=0}^{r} j\binom{r}{j}}{2^r \Phi^{-1}(q)\phi(\Phi^{-1}(q))}.$$

It follows that the asymptotic variance of the $QPV_N^r(q)$ for any r in (3.15) is minimized for $q = 0.9308$. For instance, when $r = 2$ we obtain the asymptotic variance of $QPV_N^2(0.9308)$ equal to around $3.07\sigma^4$. The variance of the likelihood-based estimator of the sample variance equals $2\sigma^4$, thus $QPV_N^2(0.9308)$ has relative efficiency of around 0.65. With the optimal, minimum variance choice of q, the estimator retains robustness to almost 7% of outliers in each tail (at least asymptotically, as will be seen in Section 3.2.5).

3.2.4 Improving efficiency by combining sub-estimators

As already suggested by Mosteller (1946), the efficiency of the quantile-based estimator can be improved by using more pairs of quantiles. We select a p-vector of quantiles $\boldsymbol{q} = [q_1, \ldots, q_p]'$ with $1 > q_1 > \ldots > q_p > 1/2$. For each symmetric pair of quantiles, i.e. q_i and $1 - q_i$, we can compute the $QPV_N^r(q_i)$ as in (3.14) that can be combined into the estimator using an averaging scheme. The quantile-based measurement of power variation based on multiple pairs is given by

$$QPV_N^r(\boldsymbol{q}, \boldsymbol{\lambda}, p) = \sum_{i=1}^{p} \lambda_i\, QPV_N^r(q_i), \tag{3.16}$$

where $\lambda_i \in (0,1)$ is a weight assigned to the ith sub-estimator, such that $\sum_{i=1}^{p} \lambda_i = 1$. The use of multiple quantiles leads to improvement of the efficiency of the estimator. The choice of p-vector of weights $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_p]'$ and the corresponding vector of quantiles $\boldsymbol{q} = [q_1, \ldots, q_p]'$ can be optimized such that a minimum variance estimator $QPV_N^r(\hat{\boldsymbol{q}}, \hat{\boldsymbol{\lambda}}, p)$ is obtained. The quantile-based estimator using multiple pairs is constructed by using a function of multiple quantiles, thus similarly the limiting distribution is obtained by a direct application of the results from Section 3.2.2.

Proposition 1. *Suppose Assumption 1 is satisfied. Consider a p-vector of quantiles* $\boldsymbol{q} = [q_1, \ldots, q_p]'$ *with* $1 > q_1 > \ldots > q_p > 1/2$. *Combine each quantile* q_i *with its symmetric pair* $1 - q_i$, *and collect these into the* $2p$*-vector* $\vec{\boldsymbol{q}}(\boldsymbol{q}) = [\vec{q}_1, \ldots, \vec{q}_{2p}]' = [1 - q_1, 1 - q_2, \ldots, 1 - q_p, q_p, \ldots, q_2, q_1]'$ *with all quantiles in ascending order. It follows that*

$$N^{1/2}\left(QPV_N^r(\boldsymbol{q}, \boldsymbol{\lambda}, p) - \sigma^r\right) \xrightarrow{d} \mathcal{N}\Big(0, \sigma^{2r} \boldsymbol{\lambda}'\, \boldsymbol{g}'\boldsymbol{G}\, \boldsymbol{g}\, \boldsymbol{\lambda}\Big),$$

where $\boldsymbol{G}$ *is the* $2p \times 2p$ *asymptotic variance-covariance matrix of sample quantiles with* ij*th element given by*

$$G_{ij} = \big(\boldsymbol{G}(\vec{\boldsymbol{q}})\big)_{ij} = \frac{\vec{q}_i(1-\vec{q}_j)}{\phi(\Phi^{-1}(\vec{q}_i))\phi(\Phi^{-1}(\vec{q}_j))}, \qquad \textit{for} \quad i \leq j; \;\; i,j = 1,\ldots,2p, \tag{3.17}$$

and where the $2p \times p$ *matrix* $\boldsymbol{g}$ *collects the partial derivatives and is given by*

$$\boldsymbol{g} = \boldsymbol{g}(\vec{\boldsymbol{q}}, r) = [\,\widetilde{\boldsymbol{g}},\, \widetilde{\boldsymbol{g}}\,\boldsymbol{J}\,]', \tag{3.18}$$

with $\widetilde{\boldsymbol{g}}$ *the* $p \times p$ *diagonal matrix with* ii*th element*

$$\widetilde{g}_{ii} = \big(\widetilde{\boldsymbol{g}}(\vec{\boldsymbol{q}}, r)\big)_{ii} = \frac{1}{c^r(\vec{q}_i)}\bigg[-\sum_{j=0}^{r} j\binom{r}{j}\big[\Phi^{-1}(\vec{q}_i)\big]^{r-1}\bigg] = \frac{1}{2^r\Phi^{-1}(\vec{q}_i)}\bigg[-\sum_{j=0}^{r} j\binom{r}{j}\bigg],$$

and the $p \times p$ $\boldsymbol{J}$ *matrix with*

$$(\boldsymbol{J})_{ij} = \begin{cases} -1 & \textit{if} \;\; j = p-i+1; \\ \;\;0 & \textit{if} \;\; j \neq p-i+1, \end{cases} \qquad \textit{for} \;\; i,j = 1,\ldots,p.$$

Proof. See Appendix 3.A.1.

Table 3.1 reports optimal values of weights $\boldsymbol{\lambda}$ and quantiles $\boldsymbol{q}$ for the number of pairs up to $p = 15$. For instance, for $p = 2$ the optimal estimator combines two sub-estimators (3.14) constructed on symmetric pairs of quantiles $(0.9770; 0.023)$ and $(0.8729; 0.1271)$, combined as in (3.16) with weights 0.4605 and 0.5395 respectively. We note that the optimal values of $\boldsymbol{\lambda}$ and $\boldsymbol{q}$ for $r = 1$ and $p = 3$ were first obtained by Ogawa (1951), and later extended by Eisenberger and Posner (1965) to include up to $p = 10$ pairs. The choice of the power r does not influence the optimal (in the minimum variance sense) values of $\boldsymbol{q}$ and $\boldsymbol{\lambda}$. It is seen from Table 3.1 that for the minimum variance estimator, the robustness property deteriorates as the optimal value of q_1 increases to unity when the number of pairs p increases. This is intuitive, as there is more information about the scale in the tails than in the central part of the distribution. Hence, a gain in efficiency in this case leads to the drawback of a lower degree of robustness. Therefore we investigate the effect of small deviations from the optimal values of $\boldsymbol{q}$ on the efficiency of the estimators.

Panel *i*) of Figure 3.2 presents the efficiency of $QPV_N^2(\boldsymbol{q}, \boldsymbol{\lambda}, p)$ relative to the maximum likelihood-based estimator of σ^2 for a number of pairs up to $p = 15$. It is seen that for one pair the efficiency of the QPV_N^2 estimator is around 65% of that of the ML estimator. As we combine more sub-estimators, the value of the minimized variance approaches the lower bound of the likelihood-based estimator rapidly and consequently the efficiency approaches unity. Panel *ii*) plots the variance of $QPV_N^2(\boldsymbol{q}, \boldsymbol{\lambda}, p)$ as a function of q_1, with $q_2, \ldots, q_p$ fixed at their optimal values, for $p \in \{2, 3, 4, 5\}$. It is seen that for q_1 declining, the variance

Table 3.1: *Optimal values of* $\boldsymbol{q}$ *and* $\boldsymbol{\lambda}$ *for the mimium variance estimator of* σ^r

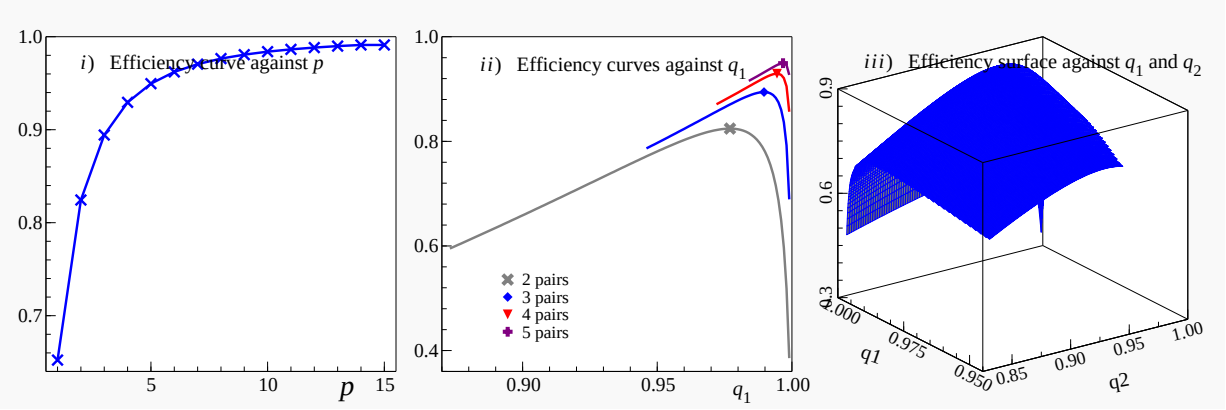
	p=1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
q_1	0.9308	0.9770	0.9896	0.9945	0.9967	0.9979	0.9986	0.9990	0.9992	0.9994	0.9996	0.9997	0.9997	0.9998	0.9998
q_2	-	0.8729	0.9452	0.9713	0.9831	0.9892	0.9927	0.9948	0.9962	0.9971	0.9978	0.9982	0.9986	0.9988	0.9990
q_3	-	-	0.8304	0.9149	0.9508	0.9689	0.9791	0.9852	0.9892	0.9919	0.9937	0.9950	0.9960	0.9968	0.9973
q_4	-	-	-	0.7983	0.8880	0.9305	0.9537	0.9676	0.9764	0.9823	0.9863	0.9892	0.9914	0.9930	0.9942
q_5	-	-	-	-	0.7731	0.8646	0.9114	0.9385	0.9555	0.9667	0.9744	0.9799	0.9840	0.9870	0.9893
q_6	-	-	-	-	-	0.7528	0.8442	0.8937	0.9238	0.9434	0.9567	0.9662	0.9730	0.9781	0.9820
q_7	-	-	-	-	-	-	0.7360	0.8263	0.8776	0.9099	0.9315	0.9467	0.9576	0.9657	0.9719
q_8	-	-	-	-	-	-	-	0.7218	0.8106	0.8628	0.8967	0.9201	0.9367	0.9490	0.9583
q_9	-	-	-	-	-	-	-	-	0.7097	0.7967	0.8493	0.8844	0.9091	0.9270	0.9405
q_{10}	-	-	-	-	-	-	-	-	-	0.6992	0.7842	0.8370	0.8729	0.8986	0.9176
q_{11}	-	-	-	-	-	-	-	-	-	-	0.6900	0.7730	0.8256	0.8621	0.8886
q_{12}	-	-	-	-	-	-	-	-	-	-	-	0.6818	0.7629	0.8152	0.8520
q_{13}	-	-	-	-	-	-	-	-	-	-	-	-	0.6745	0.7537	0.8055
q_{14}	-	-	-	-	-	-	-	-	-	-	-	-	-	0.6679	0.7452
q_{15}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.6619
λ_1	1.0000	0.4605	0.2541	0.1566	0.1040	0.0730	0.0533	0.0403	0.0312	0.0248	0.0200	0.0164	0.0136	0.0115	0.0097
λ_2	-	0.5395	0.3979	0.2771	0.1973	0.1448	0.1093	0.0846	0.0669	0.0539	0.0440	0.0365	0.0306	0.0260	0.0222
λ_3	-	-	0.3480	0.3203	0.2559	0.2004	0.1578	0.1258	0.1016	0.0832	0.0690	0.0579	0.0490	0.0419	0.0361
λ_4	-	-	-	0.2460	0.2584	0.2264	0.1897	0.1574	0.1308	0.1094	0.0922	0.0784	0.0671	0.0579	0.0503
λ_5	-	-	-	-	0.1843	0.2116	0.1980	0.1746	0.1510	0.1298	0.1116	0.0963	0.0835	0.0728	0.0638
λ_6	-	-	-	-	-	0.1439	0.1761	0.1732	0.1590	0.1420	0.1255	0.1105	0.0973	0.0858	0.0759
λ_7	-	-	-	-	-	-	0.1158	0.1487	0.1521	0.1442	0.1323	0.1196	0.1074	0.0962	0.0862
λ_8	-	-	-	-	-	-	-	0.0954	0.1272	0.1343	0.1307	0.1227	0.1131	0.1032	0.0939
λ_9	-	-	-	-	-	-	-	-	0.0801	0.1101	0.1193	0.1187	0.1135	0.1064	0.0986
λ_{10}	-	-	-	-	-	-	-	-	-	0.0683	0.0963	0.1066	0.1081	0.1051	0.0999
λ_{11}	-	-	-	-	-	-	-	-	-	-	0.0590	0.0850	0.0958	0.0987	0.0973
λ_{12}	-	-	-	-	-	-	-	-	-	-	-	0.0516	0.0755	0.0865	0.0904
λ_{13}	-	-	-	-	-	-	-	-	-	-	-	-	0.0455	0.0676	0.0785
λ_{14}	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0405	0.0609
λ_{15}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0362

curve is relatively flat near the maximum efficiency point. This is very attractive from a practical viewpoint, as it means that lowering q_1 by e.g. 2%-5% leads to only a moderate loss of the efficiency, while augmenting the robustness. This property also extends to lower quantiles. In panel *iii*) we present the efficiency surface of the estimator combining $p = 3$ sub-estimators as a function of q_1 and q_2, with q_3 fixed at its optimal value. Similarly, the efficiency surface is reasonably flat near the maximum value as q_1 and/or q_2 decrease. Panels *ii*) and *iii*) therefore nicely illustrate that gaining additional robustness against a larger fraction of outlying observations can be obtained at a relatively small drop in efficiency of the quantile-based estimator.

3.2.5 Finite sample corrections

The quantile-based estimators (3.13) and (3.14) depend on the scaling term $c(q) = \Phi^{-1}(q) - \Phi^{-1}(1-q)$, to adapt the empirical distance between quantiles for the expected distance according to the standard normal distribution. For small samples it is preferable to replace the asymptotic quantiles $\Phi^{-1}(q) \equiv Q^{\mathcal{N}}(q)$ by a term determined by the expected value of

Figure 3.2: *Efficiency of* $QPV_N^2(\boldsymbol{q}, \boldsymbol{\lambda}, p)$



the quantile for the standard normal distribution. Using quantile definition (3.6), this would deliver

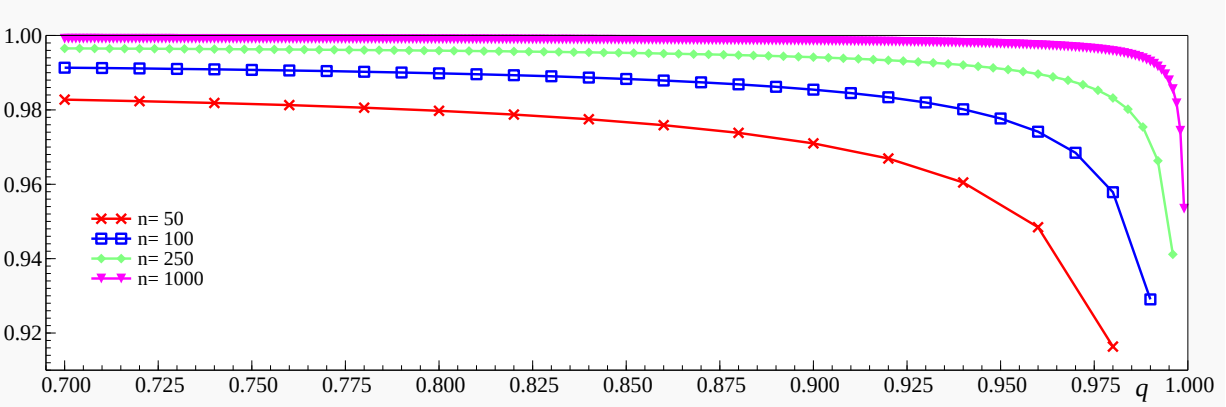
$$\overline{Q}_N^{\mathcal{N}}(q) = w\bar{Z}_{(l)} + (1-w)\bar{Z}_{(l+1)}, \tag{3.19}$$

with l and w as before, and where $\bar{Z}_{(i)}$ denotes the expected value of the ith order statistic of the independent standard normal variables $z_1, \ldots, z_N$. For a given finite N, $\bar{Z}_{(i)}$ can be obtained as

$$\bar{Z}_{(i)} = i\binom{N}{i}\int_{-\infty}^{\infty} z\big[1-\Phi(z)\big]^{N-i}\big[\Phi(z)\big]^{i-1}\phi(z)\mathrm{d}z.$$

Here ϕ and Φ are the *pdf* and *cdf* respectively of the standard normal variable, see Harter (1961).

Figure 3.3: *Finite sample correction* $\Phi^{-1}(q)/\overline{Q}_N^{\mathcal{N}}(q)$



When the number of available observations N is finite and relatively small, there are noticeable differences between the expected and limiting values of the quantiles used within

the scaling term. This is illustrated in 3.3 where the ratio between the asymptotic and final sample terms $\Phi^{-1}(q)/\overline{Q}_N^{\mathcal{N}}(q)$ are given for $N \in \{50, 100, 250, 1000\}$ and for values of $q \in \{0.7, 1\}$. Note that the difference between the asymptotic and finite sample scaling term is especially important at the more extreme quantiles, and for sample sizes less than or equal to $N = 100$. The use of asymptotic scaling term in small samples induces therefore a substantial bias, even if Assumption 1 is satisfied. In the remainder of the paper, for sample sizes $N \leq 1000$ we apply a scaling term $c_N(q)$ based on the expected quantiles $\overline{Q}_N^{\mathcal{N}}(q)$ as defined in (3.19).

The asymptotically unbiased estimator of σ^r constructed from a linear combination of sub-estimators as in (3.14) relies on the assumption of the normal distribution. However, if the assumption of the normal distribution does not hold as a fraction of outliers is found in the data, then the ordering of the normal order statistics is altered, which in turn can lead to a bias. If we know the fraction of outliers, then we are able to bias-correct the estimator by suitably adopting the scaling term $c_N(q)$ further. We investigate this issue in the simulation study in Section 3.4.

3.3 Derivation of the Outlier Test

In Section 3.3.1, we first provide the joint asymptotic distribution of the quantile- and moment-based measurements of σ^r. This joint distribution of the two estimators facilitates testing for a possible presence of outlying observations in the data. The testing procedure can be carried out through the three test statistics that we propose in Section 3.3.2.

3.3.1 Joint asymptotics of the two measurements of variation

Starting at the moment- and quantile-based measures of σ^r, it is natural to derive their joint distribution with the aim to develop formal test statistics. Under normality of the underlying observations, the resulting joint limiting distribution is shown to be bivariate normal. The joint distribution facilitates useful practical diagnostics based on these two estimators. DasGupta and Haff (2006) were the first to present the joint asymptotic distribution of the inter-quartile range and standard deviation for normally, uniformly and exponentially distributed underlying observations. Their focus is on the correlation value between these two measures, both for fixed samples and asymptotically, as both inter-quartile range and standard deviation are two well accepted measurements of data dispersion. Based on the joint asymptotic distribution DasGupta and Haff (2006, Corollary 3) provide also a thumb rule for a diagnostic to see whether the underlying data are normally distributed. Here however we focus on different power variations of the data measured by our two, robust and non-robust, estimators.

Proposition 2. *Suppose Assumption 1 is satisfied. Let $QPV_N^r(\boldsymbol{q}, \boldsymbol{\lambda}, p)$ be the quantile-based estimator of σ^r as in (3.16) and let MPV_N^r be the moment-based estimator of σ^r as in (3.1). Then as $N \to \infty$, we have*

$$N^{1/2}\Big((QPV_N^r(\boldsymbol{q}, \boldsymbol{\lambda}, p), MPV_N^r)' - (\sigma^r, \sigma^r)'\Big) \xrightarrow{d} \mathcal{N}\big(0,\, \sigma^{2r}\, \boldsymbol{\Lambda}'\, \boldsymbol{h}'\boldsymbol{H}\, \boldsymbol{h}\, \boldsymbol{\Lambda}\big),$$

where $\boldsymbol{\Lambda}$ is a $(p+1)\times 2$ matrix determined by $\boldsymbol{\lambda}$, $\boldsymbol{h}$ is a $(2p+1)\times(p+1)$ matrix collecting derivatives and $\boldsymbol{H}$ the $(2p+1)\times(2p+1)$ variance-covariance matrix of sample quantiles and sample absolute moment, given by

$$\boldsymbol{\Lambda} = \boldsymbol{\Lambda}(\boldsymbol{\lambda}) = \begin{bmatrix} \boldsymbol{\lambda}' & 0 \\ \mathbf{0}_{1\times p} & 1 \end{bmatrix}', \quad \boldsymbol{h} = \boldsymbol{h}(\vec{\boldsymbol{q}}, r) = \begin{bmatrix} \boldsymbol{g}' & 0 \\ \mathbf{0}_{1\times 2p} & \frac{1}{b(r)} \end{bmatrix}', \quad \boldsymbol{H} = \boldsymbol{H}(\vec{\boldsymbol{q}}, r) = \begin{bmatrix} \boldsymbol{G} & \boldsymbol{C} \\ \boldsymbol{C}' & M(r) \end{bmatrix}.$$

The matrices $\boldsymbol{g}$ and G are defined in (3.18), (3.17). The scalars $b(r), M(r)$ and M_z^n are specified in (3.2) *and* (3.3)*, respectively. $\boldsymbol{C}$ is the $2p$-vector that represents covariances of sample quantiles and sample absolute moment with elements*

$$C_i = ((\boldsymbol{C}(\vec{\boldsymbol{q}}, r))_i = \frac{\vec{q}_i}{\phi(\Phi^{-1}(\vec{q}_i))} M_z^r - \frac{1}{\phi(\Phi^{-1}(\vec{q}_i))} \int_{-\infty}^{\Phi^{-1}(\vec{q}_i)} |z|^r \phi(z)\, dz.$$

Proof. See Appendix 3.A.1.

The result in Proposition 2 allows us to construct test statistics and their limiting distributions under the null of normality.

3.3.2 Test statistics

The joint distribution theory developed for $QPV_N^r(\boldsymbol{q}, \boldsymbol{\lambda}, p)$ and MPV_N^r allows the construction of a formal test statistic designed to detect abnormal observations. Our tests build on the insight that under Assumption 1 the underlying observations are normally distributed, therefore both estimators measure the rth power variation of the data. In the presence of a fraction of outlying observations, the $QPV_N^r(\boldsymbol{q}, \boldsymbol{\lambda}, p)$ estimator retains robustness, provided a suitable choice of $\boldsymbol{q}$ is made such that quantiles are sufficiently far away from the extreme tails, while the MPV_N^r estimator measures a combination of the variation attributable to the standard and outlying observations. As a corollary to Proposition 2, we provide three versions of our test statistic in the spirit of Barndorff-Nielsen and Shephard (2006).

Corollary 3.3.1. *Under the setting of Proposition 2, we obtain the following test statistics:*

i) the linear test (linear BJ(p, r)):

$$\frac{N^{1/2}\Big(QPV_N^r(\boldsymbol{q}, \boldsymbol{\lambda}, p) - MPV_N^r\Big)}{\sqrt{\sigma^{2r}\Omega^{QM}}} \xrightarrow{d} \mathcal{N}\big(0,\, 1\big),$$

ii) the logarithmic test (logarithmic BJ(p, r)):

$$\frac{N^{1/2}\Big(ln\,QPV_N^r(\boldsymbol{q},\boldsymbol{\lambda},p) - ln\,MPV_N^r\Big)}{\sqrt{\Omega^{QM}}} \xrightarrow{d} \mathcal{N}\big(0,\,1\big),$$

iii) the ratio test (ratio BJ(p, r)):

$$\frac{N^{1/2}\Big(\big[QPV_N^r(\boldsymbol{q},\boldsymbol{\lambda},p)/MPV_N^r\big] - 1\Big)}{\sqrt{\Omega^{QM}}} \xrightarrow{d} \mathcal{N}\big(0,\,1\big),$$

where $\Omega^{QM} = \boldsymbol{\psi}'\,\boldsymbol{\Lambda}'\,\boldsymbol{h}'\boldsymbol{H}\,\boldsymbol{h}\,\boldsymbol{\Lambda}\,\boldsymbol{\psi}$ *and* $\boldsymbol{\psi} = [1\ -1]'$.

The linear BJ test depends on the unknown quantity σ^{2r}, therefore a feasible test needs to rely on an estimate of this quantity. Note that the possible gain in power by using a higher order r might be counteracted by a deterioration in the asymptotic variance of the test statistic, due to the estimation of σ^{2r}. In order to ensure that the linear test retains power when outliers are present in the data under the alternative, we estimate σ^{2r} with the robust quantile-based estimator, i.e. $\widehat{\sigma}^{2r} = QPV_N^{2r}(\boldsymbol{q},\boldsymbol{\lambda},p)$.

The test statistics given in Corollary 3.3.1 effectively test for the validity of the normality assumption, with the power of the test concentrated in the direction of testing for outlying observations, in any moment of the data. In the financial context, these statistics allow testing for jumps as shown in Sections 3.4 and 3.5. Two related families of tests were proposed by Barndorff-Nielsen and Shephard (2006, BNS) and Jiang and Oomen (2008, JO). We use the acronym BNS for the tests of the former authors, JO for comparable tests by Jiang and Oomen (2008), and reserve the initials BJ for the tests presented here. The BNS and JO tests are reviewed in Appendix 3.A.2. The test of BNS relates realized variance and realized bipower variation, where the former measures the quadratic variation (integrated variance plus jump component), while the latter is robust to jumps and hence measures the integrated variance. As such, the BNS test exploits the second moment of the data. The JO test relates swap variance to realized variance. Jiang and Oomen (2008) show that their test primarily uses the third moment of financial returns. Our simulation study as well as the empirical exercise will compare the performance of our tests relative to the two alternatives.

3.4 Simulation Results

In this part, we study the finite sample properties of the test statistics given in Corollary 3.3.1. We first describe the design of our simulation study. Then we discuss the performance of the quantile-based estimator in terms of the bias in small samples. We finally discuss the behavior of the testing procedure and relate it to results for the alternative approaches.

3.4.1 Simulation design

Our empirical illustration in Section 3.5 studies financial data to test for the presence of jumps. Therefore, the simulation study is set up around the jump diffusion model. The log price process X_t is assumed to follow a Brownian semi-martingale plus jumps defined as

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t + J_t \mathrm{d}C_t, \qquad t \in [0,1], \tag{3.20}$$

where μ_t denotes the drift coefficient, $\sigma_t > 0$ the spot volatility, W_t is a standard Brownian motion, J_t is a random-sized jump with mean $\mu_J(t)$ and variance $\sigma_J^2(t)$ and C_t is a process that counts the finite number of jumps in the price path up to time t. Price observations are assumed to be available at normalized equidistant times $0 = t_0 < t_1 < \ldots < t_N = 1$, where the interval $[0,1]$ denotes a trading day. Since our theory applies only when the variance is constant, we assume $\sigma_t = \sigma = 1$ which corresponds to around 16% annualized volatility and is a realistic level of financial volatility. The constant volatility model with jumps was proposed by Merton (1980) who stressed necesessity to distinguish normal and abnormal price changes. The constant variance assumption is not necessarily restrictive and can be considered when data is sampled over a short horizon or on a suitably deformed time scale.

To simulate the price process in (3.20) we use an Euler discretization scheme. For simplicity we assume $\mu_t = 0$. We generate one day of data at the highest frequency considered, setting $\Delta t = \frac{1}{N_{\max}}$. From each generated price path we sample price observations at different frequencies as determined by the choice of N. We consider $N \in \{50, 250, 1000, 5000\}$, to have situations of relatively low and high frequency. For each simulation, we use 100 000 replications to lower simulation variance. To keep output manageable, we report results for $p \in \{1,2,5\}$ and $r \in \{1,2,3,4,6\}$. The nominal size is set at 5%.

To examine the bias of the quantile-based measure and to study the power of our tests, we simulate the price process in (3.20) by adding random-sized jumps. The number of jumps is set to be either one or three per day. The arrival times of the jumps are uniformly distributed over the day, and their sizes are drawn from a normal distribution with mean $\mu_J(t) = 0$ and standard deviation chosen as a multiple κ of the daily standard deviation σ, in particular $\sigma_J(t) = \sigma_J = \kappa\sigma$ with $\kappa \in \{0, 0.25, 0.50, 1.00\}$. The choice of κ determines the relative contribution of jumps to the price variation, and here it varies from almost 6% (one jump with $\kappa = 0.25$) to 75% (three jumps with $\kappa = 1.00$).[1] The ability to detect jumps in asset prices naturally deteriorates with lower sampling frequency. This effect will be apparent in our simulation results as well, especially for combinations of both low κ and N.

[1]The relative jump contribution is measured by $100 \times \big(\sum_{i=1}^{C_1} J_i^2\big) / \big(\int_0^1 \sigma_t^2 dt + \sum_{i=1}^{C_1} J_i^2\big)$.

3.4.2 Bias of the estimator

We start by studying the small sample performance of the quantile-based estimator of powers of volatility. As we already discussed in Section 3.2.5, the quantile-based estimator can experience a minor bias when jumps are added to the price process because the ordering of the normal order statistics (as under the null) is distorted. As a result, empirical quantiles used to construct the sub-estimator (3.14) might be biased. In general, the bias depends on the fraction of jumps and their variance.

Table 3.2 presents the simulation results for this setup. For the smallest sample size considered, $N = 50$, the upward bias can be rather substantial as the first quantile q_1 may be too extreme. In particular, in case we generate three jumps, then for all quantile measures we observe a substantial bias increasing rapidly with larger size of jumps κ. Since the number of jumps in general is not known a-priori, and the jumps are randomly-sized, the changes in ordering are stochastic and no obvious bias correction is feasible.

When we increase the sample size to $N = 250$, the resulting bias diminishes, apart from the estimator combining $p = 5$ pairs which makes use of yet too extreme quantiles to retain robustness in this sample size. The impact of jumps at this sampling frequency increases relative to sample size determined by $N = 50$, the bias thus can be remarkably larger. In general, the gain in efficiency by combining more pairs of quantiles can only be obtained at the cost of lowering robustness as more and more information is extracted from the tails. In such a case, one could always lower the value of q_1 as discussed in Section 3.2.3 and observe the behavior of the measure.

Increasing the sample size further to $N = 1000$, the estimator using $p = 5$ pairs retains robustness for the one jump case, while it still makes use of the $3rd$ and $998th$ order statistics, which may be affected by a jump if three jumps are randomly placed onto the price process. Consequently, we may still observe considerable bias in this scenario. For $N = 5000$, the bias is negligible in this setup, with no estimates deviating more than 2.5% from the true value of the volatility.

3.4.3 Size of the test: Distribution under the *null*

To study the size of our test statistics, we simulate the price process as discussed in Section 3.4.1 under the null of no jumps, i.e. with $C_1 = 0$ (or $\kappa = 0$, which of course gives the same data generating process). Table 3.3 reports the sample statistics: standard deviation, skewness, kurtosis and empirical size of the proposed test statistics. We report analogous results for the BNS tests, and skip results for the JO tests as those are qualitatively similar.

The results in Table 3.3 show that in a small sample of $N = 50$ the densities of the BJ test statistics with higher order r are asymmetric and have heavier tails than the standard normal density as judged by values of skewness and kurtosis respectively. In particular, we

Table 3.2: *Bias of* $QPV_N^r(\boldsymbol{q}, \boldsymbol{\lambda}, p)$

		$C_1 = 1$ jump					$C_1 = 3$ jumps				
		$r = 1$	2	3	4	6	1	2	3	4	6
$N = 50$											
κ=0	p=1	-0.016	1.351	4.135	8.422	22.194	-0.016	1.351	4.135	8.422	22.194
	p=2	-0.020	1.656	5.088	10.426	27.947	-0.020	1.656	5.088	10.426	27.947
	p=5	*-0.012*	*1.777*	*5.440*	*11.153*	*30.079*	*-0.012*	*1.777*	*5.440*	*11.153*	*30.079*
κ=0.25	p=1	1.896	5.307	10.355	17.236	37.727	*5.922*	*13.891*	*24.289*	*37.635*	*76.217*
	p=2	*3.404*	*9.316*	*18.310*	*31.325*	*76.376*	*9.859*	*24.463*	*45.929*	*77.825*	*202.360*
	p=5	*3.099*	*8.727*	*17.464*	*30.296*	*75.984*	*9.096*	*22.726*	*42.944*	*73.341*	*196.391*
κ=0.5	p=1	3.398	8.458	15.384	24.469	50.873	*11.653*	*26.870*	*46.736*	*72.785*	*153.203*
	p=2	*12.282*	*33.940*	*72.746*	*145.304*	*588.317*	*33.161*	*98.309*	*236.120*	*550.658*	*3355.952*
	p=5	*10.345*	*29.263*	*64.571*	*134.357*	*621.139*	*29.167*	*86.524*	*210.586*	*507.160*	*3524.887*
κ=1	p=1	4.444	10.664	18.919	29.574	60.219	*18.399*	*44.247*	*81.785*	*138.737*	*385.441*
	p=2	*34.686*	*121.353*	*360.713*	*1090.803*	*12320.874*	*88.213*	*384.392*	*1510.529*	*6256.900*	*136750.650*
	p=5	*27.855*	*102.074*	*326.327*	*1089.530*	*15762.049*	*76.022*	*331.723*	*1352.230*	*6015.149*	*161986.657*
$N = 250$											
κ=0	p=1	0.003	0.306	0.910	1.819	4.586	0.003	0.306	0.910	1.819	4.586
	p=2	0.004	0.353	1.051	2.104	5.319	0.004	0.353	1.051	2.104	5.319
	p=5	*-0.001*	*0.429*	*1.295*	*2.614*	*6.708*	*-0.001*	*0.429*	*1.295*	*2.614*	*6.708*
κ=0.25	p=1	0.612	1.533	2.769	4.332	8.491	1.850	4.050	6.621	9.586	16.814
	p=2	0.801	1.967	3.511	5.450	10.596	2.483	5.428	8.874	12.864	22.698
	p=5	*2.877*	*7.738*	*16.103*	*31.231*	*120.673*	*7.893*	*21.866*	*48.993*	*106.890*	*582.305*
κ=0.5	p=1	0.756	1.824	3.211	4.931	9.426	2.302	4.977	8.049	11.547	19.967
	p=2	1.030	2.433	4.224	6.423	12.146	3.270	7.077	11.473	16.521	28.883
	p=5	*8.150*	*28.088*	*83.005*	*252.323*	*2955.075*	*20.569*	*84.231*	*316.398*	*1272.751*	*26912.086*
κ=1	p=1	0.834	1.981	3.450	5.254	9.930	2.552	5.490	8.841	12.639	21.729
	p=2	1.154	2.685	4.610	6.951	12.987	3.724	8.034	12.991	18.672	32.574
	p=5	*19.501*	*100.611*	*503.176*	*2794.662*	*118683.742*	*46.490*	*326.882*	*2346.692*	*18760.305*	*1594426.593*
$N = 1000$											
κ=0	p=1	-0.012	0.053	0.194	0.412	1.080	-0.012	0.053	0.194	0.412	1.080
	p=2	-0.010	0.068	0.234	0.489	1.269	-0.010	0.068	0.234	0.489	1.269
	p=5	-0.007	0.094	0.302	0.619	1.583	-0.007	0.094	0.302	0.619	1.583
κ=0.25	p=1	0.171	0.419	0.744	1.148	2.195	0.539	1.159	1.861	2.647	4.476
	p=2	0.229	0.549	0.959	1.460	2.744	0.711	1.519	2.427	3.436	5.773
	p=5	0.383	0.882	1.500	2.240	4.103	*1.481*	*3.279*	*5.510*	*8.374*	*17.809*
κ=0.5	p=1	0.189	0.455	0.799	1.222	2.307	0.596	1.274	2.035	2.880	4.833
	p=2	0.257	0.604	1.042	1.572	2.914	0.798	1.695	2.694	3.797	6.329
	p=5	0.442	1.003	1.685	2.491	4.498	*2.381*	*6.056*	*12.641*	*26.548*	*154.305*
κ=1	p=1	0.199	0.474	0.828	1.260	2.365	0.625	1.333	2.123	2.999	5.015
	p=2	0.271	0.633	1.085	1.630	3.001	0.843	1.786	2.831	3.983	6.615
	p=5	0.473	1.066	1.781	2.623	4.705	*4.104*	*14.368*	*48.562*	*188.396*	*4241.319*
$N = 5000$											
κ=0	p=1	0.026	0.067	0.124	0.196	0.386	0.026	0.067	0.124	0.196	0.386
	p=2	0.041	0.099	0.175	0.268	0.509	0.041	0.099	0.175	0.268	0.509
	p=5	0.069	0.160	0.272	0.406	0.740	0.069	0.160	0.272	0.406	0.740
κ=0.25	p=1	0.066	0.148	0.245	0.357	0.629	0.147	0.310	0.489	0.684	1.120
	p=2	0.093	0.205	0.334	0.481	0.828	0.199	0.417	0.653	0.908	1.473
	p=5	0.149	0.319	0.513	0.728	1.228	0.312	0.648	1.008	1.394	2.242
κ=0.5	p=1	0.068	0.151	0.250	0.364	0.640	0.152	0.320	0.504	0.704	1.150
	p=2	0.096	0.209	0.341	0.490	0.843	0.207	0.432	0.676	0.938	1.518
	p=5	0.153	0.328	0.526	0.746	1.255	0.326	0.676	1.051	1.452	2.331
κ=1	p=1	0.069	0.153	0.253	0.368	0.645	0.155	0.326	0.512	0.714	1.166
	p=2	0.097	0.212	0.345	0.495	0.851	0.211	0.440	0.687	0.953	1.541
	p=5	0.155	0.333	0.533	0.756	1.270	0.333	0.691	1.073	1.482	2.377

Note: The table reports the percentage bias of σ^r averaged over number of repetitions. In italics, the biases for cases where the maximum quantile q_1 cannot be estimated robustly for lack of sufficient observations, i.e. when $q_1 > \frac{N-C_1}{N+1}$ when $\kappa > 0$, or $q_1 > \frac{N}{N+1}$ otherwise.

see a poor performance of the linear BJ version, which however is not surprising given that this statistic involves division by a sample estimate of σ^{2r}. For instance, the linear BJ$(\cdot, 6)$ test requires an estimate of σ^{12} which is a difficult quantity to estimate, particularly in small samples. Since the BJ log- and ratio tests alleviate the need to estimate σ^{2r}, they are found, as expected, to have better small sample performance than the linear test. The ratio

Table 3.3: *Distribution under the null*

		BJ(·, 1)			BJ(·, 2)			BJ(·, 3)			BJ(·, 4)			BJ(·, 6)			BNS			
		lin	log	ratio	lin	log	ratio	lin	log	ratio	lin	log	ratio	lin	log	ratio	lin	log	ratio	adj.ratio
N=50																				
std.dev.	BJ(1,·)	0.927	0.935	0.952	0.835	0.834	0.854	0.850	0.795	0.806	0.940	0.775	0.803	1.269	0.672	0.895	1.237	1.161	1.116	0.999
	BJ(2,·)	0.984	1.000	1.018	0.720	0.753	0.784	0.515	0.560	0.601	0.455	0.539	0.632	0.330	0.474	0.807	-	-	-	-
	BJ(5,·)	1.043	1.061	1.076	0.669	0.710	0.737	0.488	0.552	0.598	0.496	0.611	0.723	0.356	0.527	0.908	-	-	-	-
skew	BJ(1,·)	-0.154	0.078	0.313	-0.714	-0.282	0.108	-1.932	-0.836	-0.152	-3.753	-1.060	-0.073	-17.094	-1.042	0.481	-1.282	-0.820	-0.444	-0.286
	BJ(2,·)	-0.060	0.130	0.318	-0.242	-0.012	0.200	-0.727	-0.378	-0.062	-0.867	-0.211	0.347	-1.841	-0.066	1.256	-	-	-	-
	BJ(5,·)	-0.076	0.076	0.221	-0.153	-0.001	0.104	-0.632	-0.283	-0.019	-1.022	-0.250	0.368	-2.178	-0.209	1.368	-	-	-	-
kurt	BJ(1,·)	3.017	2.995	3.164	4.205	3.297	3.110	11.297	4.369	2.996	39.452	4.810	2.658	838.859	4.498	2.684	6.241	4.220	3.287	3.014
	BJ(2,·)	2.993	3.049	3.217	3.100	3.027	3.097	3.969	3.325	3.046	4.447	3.218	3.185	9.803	3.257	5.340	-	-	-	-
	BJ(5,·)	2.991	3.026	3.105	3.076	3.140	3.159	4.247	3.861	3.559	5.376	3.793	3.638	11.561	3.611	6.761	-	-	-	-
size	BJ(1,·)	0.034	0.038	0.044	0.020	0.021	0.030	0.025	0.015	0.021	0.033	0.014	0.023	0.037	0.006	0.094	0.088	0.080	0.075	0.048
	BJ(2,·)	0.049	0.054	0.060	0.010	0.020	0.031	0.001	0.001	0.007	0.000	0.001	0.021	0.001	0.000	0.082	-	-	-	-
	BJ(5,·)	0.065	0.072	0.077	0.034	0.052	0.065	0.001	0.011	0.026	0.001	0.013	0.054	0.001	0.004	0.110	-	-	-	-
N=250																				
std.dev.	BJ(1,·)	0.987	0.989	0.992	0.975	0.974	0.978	0.982	0.968	0.971	1.002	0.959	0.968	1.072	0.885	0.966	1.045	1.032	1.024	0.989
	BJ(2,·)	0.984	0.987	0.990	0.934	0.939	0.945	0.923	0.924	0.929	0.932	0.919	0.927	0.968	0.841	0.904	-	-	-	-
	BJ(5,·)	1.006	1.010	1.013	0.837	0.848	0.855	0.785	0.808	0.822	0.783	0.826	0.855	0.684	0.751	0.847	-	-	-	-
skew	BJ(1,·)	-0.053	0.059	0.173	-0.290	-0.095	0.099	-0.654	-0.295	0.037	-1.097	-0.412	0.131	-4.695	-0.537	0.584	-0.533	-0.370	-0.214	-0.155
	BJ(2,·)	-0.033	0.047	0.128	-0.223	-0.100	0.020	-0.667	-0.399	-0.157	-1.159	-0.540	-0.074	-4.325	-0.712	0.367	-	-	-	-
	BJ(5,·)	-0.031	0.031	0.092	-0.071	-0.014	0.040	-0.309	-0.180	-0.047	-0.485	-0.214	0.096	-1.217	-0.385	0.469	-	-	-	-
kurt	BJ(1,·)	3.031	3.041	3.096	3.172	3.041	3.045	3.812	3.155	2.976	5.867	3.308	2.937	143.154	3.603	3.373	3.607	3.304	3.112	3.035
	BJ(2,·)	2.986	2.990	3.016	3.172	3.089	3.061	4.054	3.430	3.121	6.636	3.782	3.090	93.735	4.247	3.268	-	-	-	-
	BJ(5,·)	2.996	3.004	3.020	3.034	3.030	3.031	3.170	3.119	3.054	3.400	3.210	3.084	5.963	3.536	3.447	-	-	-	-
size	BJ(1,·)	0.047	0.048	0.048	0.044	0.044	0.047	0.043	0.042	0.046	0.044	0.038	0.047	0.048	0.025	0.063	0.058	0.057	0.055	0.047
	BJ(2,·)	0.047	0.048	0.049	0.037	0.040	0.043	0.031	0.035	0.040	0.031	0.033	0.042	0.034	0.021	0.052	-	-	-	-
	BJ(5,·)	0.052	0.054	0.054	0.032	0.035	0.038	0.016	0.023	0.029	0.011	0.024	0.037	0.009	0.012	0.049	-	-	-	-
N=1000																				
std.dev.	BJ(1,·)	0.994	0.995	0.996	0.994	0.994	0.995	0.997	0.993	0.995	1.002	0.991	0.994	1.014	0.959	0.976	1.009	1.006	1.004	0.989
	BJ(2,·)	0.996	0.997	0.998	0.991	0.992	0.994	0.988	0.988	0.989	0.989	0.984	0.985	0.994	0.946	0.953	-	-	-	-
	BJ(5,·)	0.999	1.000	1.000	0.945	0.947	0.949	0.939	0.944	0.946	0.945	0.951	0.955	0.941	0.916	0.924	-	-	-	-
skewn	BJ(1,·)	-0.026	0.030	0.086	-0.137	-0.041	0.056	-0.305	-0.134	0.034	-0.492	-0.199	0.081	-1.201	-0.363	0.289	-0.253	-0.178	-0.104	-0.081
	BJ(2,·)	-0.007	0.034	0.075	-0.126	-0.063	0.000	-0.334	-0.209	-0.088	-0.536	-0.292	-0.062	-1.373	-0.511	0.103	-	-	-	-
	BJ(5,·)	-0.002	0.029	0.059	-0.085	-0.054	-0.024	-0.281	-0.201	-0.122	-0.458	-0.268	-0.084	-1.417	-0.562	0.027	-	-	-	-
kurt	BJ(1,·)	2.998	2.999	3.011	3.012	2.983	2.986	3.148	3.008	2.973	3.436	3.049	2.972	6.324	3.369	3.137	3.106	3.046	3.007	2.990
	BJ(2,·)	3.015	3.018	3.027	3.081	3.059	3.051	3.248	3.117	3.041	3.615	3.215	3.038	7.747	3.745	3.077	-	-	-	-
	BJ(5,·)	3.009	3.012	3.017	3.082	3.074	3.068	3.224	3.155	3.099	3.584	3.299	3.137	8.831	4.108	3.204	-	-	-	-
size	BJ(1,·)	0.049	0.049	0.050	0.048	0.048	0.049	0.048	0.048	0.049	0.049	0.047	0.050	0.045	0.041	0.050	0.052	0.051	0.050	0.047
	BJ(2,·)	0.049	0.050	0.050	0.049	0.049	0.050	0.046	0.048	0.049	0.044	0.046	0.049	0.040	0.038	0.046	-	-	-	-
	BJ(5,·)	0.051	0.052	0.052	0.042	0.043	0.044	0.037	0.039	0.041	0.037	0.041	0.044	0.032	0.034	0.043	-	-	-	-
N=5000																				
std.dev.	BJ(1,·)	1.002	1.003	1.003	1.000	1.000	1.001	0.999	0.999	1.000	1.000	0.998	1.000	1.002	0.991	0.994	1.000	0.999	0.999	0.993
	BJ(2,·)	1.000	1.000	1.001	0.996	0.998	0.999	0.996	0.997	0.999	0.996	0.997	0.999	0.997	0.988	0.990	-	-	-	-
	BJ(5,·)	1.003	1.004	1.005	0.987	0.989	0.991	0.985	0.988	0.991	0.986	0.990	0.994	0.986	0.981	0.984	-	-	-	-
skew	BJ(1,·)	-0.042	-0.016	0.009	-0.081	-0.037	0.008	-0.149	-0.071	0.006	-0.230	-0.100	0.029	-0.544	-0.215	0.088	-0.115	-0.081	-0.048	-0.038
	BJ(2,·)	0.004	0.022	0.040	-0.058	-0.031	-0.003	-0.151	-0.098	-0.045	-0.239	-0.137	-0.037	-0.596	-0.291	-0.015	-	-	-	-
	BJ(5,·)	-0.004	0.009	0.023	-0.038	-0.024	-0.010	-0.135	-0.099	-0.063	-0.219	-0.138	-0.057	-0.637	-0.343	-0.084	-	-	-	-
kurt	BJ(1,·)	3.016	3.013	3.014	3.034	3.024	3.021	3.068	3.032	3.019	3.128	3.040	3.013	3.834	3.187	3.052	3.033	3.022	3.015	3.010
	BJ(2,·)	2.993	2.995	2.997	2.990	2.987	2.986	3.018	2.994	2.979	3.104	3.031	2.993	4.043	3.293	3.063	-	-	-	-
	BJ(5,·)	2.975	2.975	2.976	3.024	3.022	3.021	3.077	3.061	3.047	3.153	3.091	3.052	4.418	3.463	3.144	-	-	-	-
size	BJ(1,·)	0.050	0.050	0.051	0.050	0.050	0.051	0.050	0.050	0.051	0.050	0.050	0.051	0.048	0.048	0.050	0.050	0.050	0.050	0.049
	BJ(2,·)	0.051	0.051	0.051	0.050	0.050	0.051	0.049	0.050	0.051	0.048	0.049	0.051	0.046	0.047	0.050	-	-	-	-
	BJ(5,·)	0.055	0.055	0.056	0.067	0.068	0.069	0.053	0.054	0.056	0.048	0.051	0.053	0.043	0.045	0.050	-	-	-	-

BJ test has a somewhat better behavior than the BJ log version except when r becomes large, $r = 6$ say. At this sample size, our tests with lower orders of r perform considerably better than with higher orders as the descriptive statistics that are obtained correspond closer to the asymptotic standard normal distribution.

As we increase the sample size and look at the simulation results based on $N = 250$, we observe a noticeable improvement in performance. The values of variance, skewness and kurtosis are more consistent with the standard normal distribution. The size distortions are far less severe, as the empirical sizes match the nominal size of 5% closely. At this sampling frequency the linear BJ test still performs worse than the log- and ratio BJ test statistics. Also, the ratio BJ test seems to perform slightly better than the log BJ test when judged by the reported descriptive statistics. When we increase the sample sizes to $N = 1000$ and $N = 5000$, then the empirical sizes are consistent with the nominal size of 5%.[2]

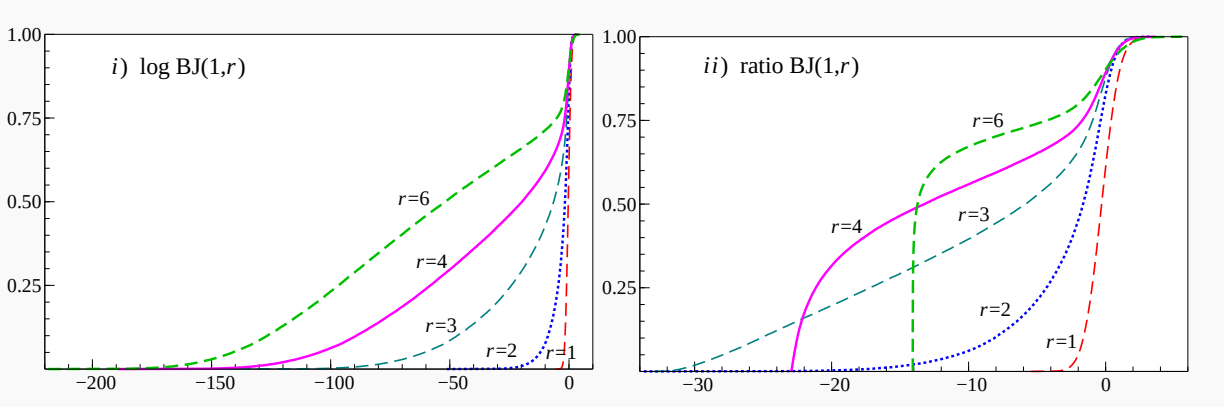
3.4.4 Power of the test: Detection of jumps

To examine the power of our tests, we simulate the price process by adding random jumps as discussed in Section 3.4.1. The construction of the test would suggest that the power of the test will increase with the order r, as the moment-based measure becomes more heavily influenced by outliers. To study this effect, Figure 3.4 presents the empirical cumulative distribution function of the simulated log (left panel) and ratio (right panel) test-statistics when a single random-sized jump with $\kappa = 0.25$ is added to $N = 5000$ observations. Our tests are one-sided in nature and their distributions become left-skewed under the alternative. The most important aspect is that the test statistics built on higher orders of r are more skewed to the left, reflecting higher impact of outliers on the MPV_N^r measure. For the ratio test statistic, note that there is a limit of $-\sqrt{N/\Omega^{QM}}$ for the test statistic under the alternative, which is attained when jumps drive $MPV^r \to \infty$, hence in general for increasing r. We examine the performance of the tests when N is considerably smaller. To conserve on space we report results exclusively based on the BJ ratio test.

Table 3.4 presents the rejection rates of our test, with analogous results reported for the (adjusted) BNS and JO ratio tests. The simulation results can be summarized as follows.

[2]The theory developed in this chapter is based on the constant volatility assumption, as the test effectively relates two measures of unconditional variation under normality. If this assumption is not met, i.e. if there is remaining deterministic or stochastic time-varying volatility, then the density of returns will have heavier tails than the normal density. Our test on such data will detect those tails, and hence it becomes oversized. A simulation study (results not reported here) indeed indicates that the size distortion is minor for approximately constant volatility processes. The size however increases up to 25% when volatility process has common GARCH persistence, or if there is periodic variance pattern in returns such as for intraday data. However, as explained in the empirical section of this chapter, we propose to devolatilize the data before applying the jumps test.

Figure 3.4: *Distribution under alternative*



The use of the BJ tests built on higher order of r does improve the power to detect jumps considerably , also for relatively large sample sizes. This is fully in line with expectations. For instance, for sample sizes $N = 1000$ or $N = 5000$ the BJ$(\cdot, 6)$ tests are uniformly more powerful than any other test considered here, regardless of the variance of jumps and of the number of jumps.

We find, as expected, that the BJ$(\cdot, 2)$ test performs similarly to the BNS test as both use the second moment of the data. However, unlike the BJ$(\cdot, 2)$ test, the BNS test can (substantially) lose power when jumps occur at adjacent locations, as a requirement for the robustness of both bipower and quadpower measures is violated in such a case. The BJ$(\cdot, 3)$ test is found to perform similarly (in the single jump case) to the JO test, which primarily uses the third order of the data; see Jiang and Oomen (2008). However, unlike the BJ$(\cdot, 3)$ test, the JO test can (substantially) lose power when multiple jumps occur with different signs as the cubed jumps then offset each other.

If we decrease the sample size to $N = 1000$, then the advantage of the BJ$(\cdot, 6)$ test over the BJ$(\cdot, 4)$ version decreases, but still both versions are more powerful than the BNS or JO test, and the difference in power can be substantial. For instance, with three jumps with variance determined by $\kappa = 0.25$, the BJ$(5, 4)$ has a power of 92.04% and the BJ$(5, 6)$ has a power of 92.42%, rejection rates considerably higher than the 77.66% of the JO test and 72.42% of the adjusted BNS. If we increase the variance of jumps to e.g. $\kappa = 1.00$, then the power of all tests is close to unity.

For the sample size $N = 250$, a sample size for which the size distortions are modest, the BJ tests based on order $r = 3$ and $r = 4$ are in general outperforming other BJ versions. Considering the case with a single jump, we find that the BJ$(\cdot, 3)$ and BJ$(\cdot, 4)$ tests perform still slightly better than the JO test and the BNS tests. When multiple jumps are considered, then superiority of the BJ test with a larger number of pairs and

Table 3.4: *Power of ratio tests*

		1 jump								3 jumps							
		BJ(·, 1)	BJ(·, 2)	BJ(·, 3)	BJ(·, 4)	BJ(·, 6)	adjBNS	BNS	JO	BJ(·, 1)	BJ(·, 2)	BJ(·, 3)	BJ(·, 4)	BJ(·, 6)	adjBNS	BNS	JO
$N = 50$																	
κ=0.25	BJ(1,·)	0.0466	0.0433	0.0418	0.0288	0.0796	0.1003	0.0687	0.1493	0.0526	0.0663	0.0751	0.0363	0.0558	0.1446	0.1062	0.2235
	BJ(2,·)	0.1025	0.0562	0.0060	0.0172	0.0685	-	-	-	0.1724	0.0972	0.0052	0.0118	0.0451	-	-	-
	BJ(5,·)	0.1177	0.0941	0.0261	0.0478	0.0881	-	-	-	0.1948	0.1480	0.0288	0.0410	0.0584	-	-	-
κ=0.50	BJ(1,·)	0.0634	0.1660	0.2014	0.1392	0.0519	0.2359	0.1970	0.3417	0.1084	0.3592	0.4365	0.2995	0.0176	0.4673	0.4174	0.5585
	BJ(2,·)	0.2855	0.1627	0.0217	0.0333	0.0437	-	-	-	0.5508	0.3324	0.0432	0.0381	0.0126	-	-	-
	BJ(5,·)	0.2911	0.1502	0.1259	0.1379	0.0566	-	-	-	0.5827	0.3335	0.2131	0.1988	0.0166	-	-	-
κ=1.00	BJ(1,·)	0.1929	0.4498	0.4957	0.4286	0.0287	0.4981	0.4646	0.6013	0.4260	0.7889	0.8420	0.7681	0.0031	0.8329	0.8062	0.8309
	BJ(2,·)	0.5602	0.2841	0.2032	0.2251	0.0247	-	-	-	0.8708	0.5300	0.2728	0.2493	0.0024	-	-	-
	BJ(5,·)	0.5577	0.1610	0.4097	0.4152	0.0317	-	-	-	0.8947	0.5044	0.5195	0.4918	0.0031	-	-	-
$N = 250$																	
κ=0.25	BJ(1,·)	0.0567	0.1600	0.2485	0.2783	0.2193	0.1742	0.1631	0.2724	0.0877	0.4011	0.5633	0.6053	0.4790	0.4086	0.3934	0.5074
	BJ(2,·)	0.0612	0.2091	0.2821	0.3011	0.2545	-	-	-	0.0962	0.4945	0.6192	0.6426	0.5528	-	-	-
	BJ(5,·)	0.2244	0.1534	0.1881	0.2587	0.2286	-	-	-	0.4794	0.3368	0.4231	0.5719	0.4412	-	-	-
κ=0.50	BJ(1,·)	0.1105	0.4350	0.5398	0.5682	0.5150	0.4447	0.4341	0.5552	0.3496	0.8369	0.9049	0.9176	0.8779	0.8299	0.8238	0.8380
	BJ(2,·)	0.1471	0.5012	0.5744	0.5899	0.5501	-	-	-	0.4122	0.8793	0.9215	0.9279	0.9021	-	-	-
	BJ(5,·)	0.5018	0.2938	0.4890	0.5550	0.5279	-	-	-	0.8487	0.6563	0.7975	0.9065	0.7726	-	-	-
κ=1.00	BJ(1,·)	0.3073	0.6887	0.7556	0.7716	0.7408	0.6932	0.6866	0.7629	0.7778	0.9730	0.9855	0.9878	0.9811	0.9707	0.9694	0.9547
	BJ(2,·)	0.3883	0.7304	0.7746	0.7837	0.7613	-	-	-	0.8307	0.9813	0.9882	0.9895	0.9854	-	-	-
	BJ(5,·)	0.7287	0.3199	0.7248	0.7644	0.7482	-	-	-	0.9718	0.8407	0.9214	0.9863	0.8617	-	-	-
$N = 1000$																	
κ=0.25	BJ(1,·)	0.0670	0.3226	0.4774	0.5355	0.5667	0.3350	0.3301	0.4674	0.1518	0.7267	0.8662	0.9016	0.9163	0.7278	0.7242	0.7766
	BJ(2,·)	0.0794	0.4064	0.5212	0.5577	0.5758	-	-	-	0.2033	0.8115	0.8948	0.9144	0.9219	-	-	-
	BJ(5,·)	0.0877	0.4839	0.5488	0.5700	0.5804	-	-	-	0.1990	0.8551	0.9090	0.9204	0.9242	-	-	-
κ=0.50	BJ(1,·)	0.1416	0.6038	0.7183	0.7531	0.7720	0.6134	0.6104	0.7092	0.4880	0.9500	0.9787	0.9850	0.9874	0.9502	0.9494	0.9451
	BJ(2,·)	0.2022	0.6677	0.7442	0.7666	0.7774	-	-	-	0.6179	0.9690	0.9841	0.9872	0.9882	-	-	-
	BJ(5,·)	0.2490	0.7229	0.7617	0.7746	0.7802	-	-	-	0.6085	0.9762	0.9862	0.9879	0.9887	-	-	-
κ=1.00	BJ(1,·)	0.3537	0.7913	0.8560	0.8748	0.8843	0.7985	0.7966	0.8502	0.8538	0.9926	0.9970	0.9978	0.9982	0.9925	0.9924	0.9870
	BJ(2,·)	0.4574	0.8280	0.8701	0.8818	0.8871	-	-	-	0.9131	0.9955	0.9977	0.9980	0.9983	-	-	-
	BJ(5,·)	0.5218	0.8582	0.8789	0.8856	0.8887	-	-	-	0.8867	0.9964	0.9979	0.9982	0.9983	-	-	-
$N = 5000$																	
κ=0.25	BJ(1,·)	0.0702	0.5075	0.6846	0.7416	0.7812	0.5320	0.5304	0.6649	0.1973	0.9081	0.9712	0.9831	0.9890	0.9176	0.9171	0.9313
	BJ(2,·)	0.0877	0.5837	0.7130	0.7542	0.7847	-	-	-	0.2884	0.9445	0.9783	0.9856	0.9894	-	-	-
	BJ(5,·)	0.1036	0.6655	0.7383	0.7647	0.7882	-	-	-	0.3593	0.9679	0.9831	0.9872	0.9900	-	-	-
κ=0.50	BJ(1,·)	0.1523	0.7355	0.8369	0.8687	0.8898	0.7501	0.7494	0.8266	0.5604	0.9861	0.9957	0.9974	0.9984	0.9873	0.9872	0.9856
	BJ(2,·)	0.2197	0.7821	0.8524	0.8756	0.8916	-	-	-	0.7059	0.9920	0.9968	0.9979	0.9984	-	-	-
	BJ(5,·)	0.2864	0.8272	0.8667	0.8813	0.8932	-	-	-	0.7848	0.9954	0.9975	0.9981	0.9984	-	-	-
κ=1.00	BJ(1,·)	0.3706	0.8650	0.9176	0.9337	0.9438	0.8725	0.8720	0.9121	0.8832	0.9979	0.9994	0.9997	0.9998	0.9982	0.9982	0.9970
	BJ(2,·)	0.4796	0.8890	0.9259	0.9372	0.9449	-	-	-	0.9384	0.9987	0.9995	0.9997	0.9998	-	-	-
	BJ(5,·)	0.5577	0.9126	0.9324	0.9396	0.9454	-	-	-	0.9597	0.9993	0.9997	0.9997	0.9998	-	-	-

with power $r = 3$ or $r = 4$ is even more evident. Since $N = 250$ is relatively small sample, then the use of more quantiles, say $p = 5$, leads to a decrease in power. For instance, with $\kappa = 0.25$ and three jumps, the power of BJ$(5, 2)$ drops to 33.68%, while the BJ$(2, 2)$ is around 49.45%. Note that with too extreme quantiles, the robustness of quantile measures is lost. Nevertheless, at this sampling frequency, our test based on specific orders of r can still outperform the two alternatives.

If we further decrease the sample size to $N = 50$, then constructing of our test on higher order of r does not improve power to detect jumps, even when additional jumps (three rather than one) are placed in the price process. Note that at this sample size, all tests actually suffer from size distortions. In conclusion, for realistic sample sizes, our test is more powerful than the BNS or JO for all jump sizes and numbers of jumps, for specific orders of r.

3.5 Empirical Applications

3.5.1 Dataset

The data we analyze includes fifteen components of the Dow Jones Industrial Average index. The data are consolidated trades extracted from the Trade and Quote (TAQ) database through the Wharton Research Data Services (WRDS) system. The sample period spans almost five years, from January 3, 2006 to September 30, 2010, with a total of $T = 1195$ trading days for all equities.

Before we construct measures and compute tests, the data needs to be cleaned. The importance of tick-by-tick data cleaning is highlighted by Brownless and Gallo (2006), Hansen and Lunde (2006) and Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009) who provide a guideline on cleaning procedures based on the TAQ qualifiers that are included in the files and described in the TAQ User's Guide available at the WRDS website. Following Barndorff-Nielsen et al. (2009) we consider five steps: **P1**. delete entries with a time stamp outside the 9:30am to 4:00pm Eastern Time window; **P2**. delete entries with transaction prices equal to zero; **P3**. retain entries only from the single (primary) exchange which is the NYSE (including NYSE Direct+) in our application; **T1**. delete incorrect trades as indicated by the correction indicator; **T2**. delete entries when the sale condition is not regular. Table 3.5 provides statistics regarding the pre-processing of the data, reporting the number of trades per year before/after cleaning, and the average number of trades per day including minimum and maximum, over the years in the sample. The cleaning procedures lead to a substantial data reduction. For all equities the largest deletion of the raw data is due to step **P3**, which is implement also to reduce the impact of time-delays in the reporting of trades updates, see Barndorff-Nielsen et al. (2009) for further discussion. The largest number of transactions for all equities was recorded over 2007 and 2008. The

Table 3.5: *Equity data: summary statistics on the number of consolidated trades*

		2006			2007			2008			2009			2010	
Symbol	Raw	Clean	Avg	Raw	Clean	Avg	Raw	Clean	Avg	Raw	Clean	Avg	Raw	Clean	Avg
AA	2 938 194	1 399 359	5 575	10 484 729	2 730 384	10 878	19 052 862	3 002 379	11 867	25 252 398	2 428 419	9 637	15 661 032	1 172 752	6 238
		(3 095; 15 103)			(3 566; 35 077)			(3 521; 25 190)			(3 844; 31 749)			(2 229; 20 235)	
AXP	2 131 247	1 149 936	4 581	6 960 172	2 466 466	9 827	18 694 570	3 171 851	12 537	22 100 137	3 507 182	13 917	9 869 786	1 671 506	8 891
		(1 849; 19 541)			(3 744; 24 526)			(4 199; 39 268)			(2 340; 43 138)			(4 133; 27 335)	
BA	2 332 928	1 297 892	5 171	6 033 728	2 326 756	9 270	10 433 805	2 581 156	10 202	9 177 067	2 030 485	8 057	5 738 717	1 431 470	7 614
		(1 837; 10 114)			(2 827; 23 193)			(3 293; 24 507)			(1 613; 19 235)			(3 126; 19 253)	
CAT	2 917 482	1 358 294	5 412	6 273 666	2 387 487	9 512	12 101 199	2 488 655	9 837	16 313 099	2 490 239	9 882	8 947 046	1 781 497	9 476
		(2 584; 11 976)			(3 027; 39 266)			(1 953; 21 282)			(2 280; 30 249)			(3 134; 26 790)	
GE	7 308 481	1 682 671	6 704	17 890 755	4 073 410	16 229	53 097 581	6 299 728	24 900	59 023 304	4 494 888	17 837	26 913 874	2 162 807	11 504
		(2 384; 25 629)			(7 257; 42 630)			(10 238; 78 211)			(2 751; 58 635)			(4 304; 42 812)	
HD	4 908 720	1 473 846	5 872	12 053 593	2 940 748	11 716	22 558 723	3 210 311	12 689	19 088 604	2 638 725	10 471	12 543 075	1 817 506	9 668
		(2 274; 21 766)			(4 413; 27 376)			(3 416; 30 495)			(2 672; 21 788)			(3 795; 33 095)	
IBM	2 811 212	1 533 389	6 109	7 417 523	3 129 273	12 467	12 485 302	3 202 219	12 657	11 580 628	2 555 039	10 139	7 042 620	1 991 539	10 593
		(2 483; 15 266)			(3 981; 31 309)			(3 874; 30 967)			(4 036; 24 791)			(4 203; 27 314)	
JPM	4 127 533	1 484 066	5 913	14 296 841	3 937 709	15 688	54 047 407	7 229 963	28 577	68 011 683	6 825 146	27 084	31 389 431	3 247 516	17 274
		(2 414; 12 739)			(5 414; 44 706)			(6 771; 89 980)			(4 561; 67 384)			(7 849; 47 420)	
KO	2 643 426	1 179 259	4 698	7 068 542	2 313 047	9 215	15 265 008	3 066 471	12 120	12 916 579	2 472 864	9 813	8 717 521	1 421 613	7 562
		(2 246; 8 131)			(2 822; 20 877)			(3 010; 35 184)			(3 192; 26 250)			(3 781; 25 352)	
MCD	2 543 832	1 169 975	4 661	6 436 406	2 028 862	8 083	12 970 858	2 680 197	10 594	12 452 641	2 298 559	9 121	6 643 969	1 389 262	7 390
		(1 830; 9 538)			(3 266; 20 252)			(2 942; 33 469)			(2 187; 23 720)			(3 156; 22560)	
PFE	9 148 652	1 627 340	6 483	14 623 994	3 235 968	12 892	27 636 290	3 477 087	13 743	32 199 521	3 298 404	13 089	20 985 874	1 447 425	7 699
		(3 022; 12 164)			(4 444; 21 838)			(5 410; 41 934)			(2 844; 45 629)			(3 658; 24 496)	
PG	3 064 038	1 624 277	6 471	8 667 109	2 863 358	11 408	19 369 125	3 658 744	14 461	16 275 148	2 803 493	11 125	10 602 195	1 656 146	8 809
		(3 404; 16 053)			(4 046; 26 284)			(6 299; 37 735)			(2 499; 24 520)			(4 237; 27 580)	
T	5 183 297	1 663 205	6 626	13 641 001	3 439 970	13 705	29 034 642	4 198 031	16 593	25 862 175	3 076 086	12 207	16 199 893	1 430 443	7 609
		(1 807; 18 575)			(5 348; 27 409)			(5 282; 45 161)			(2 334; 28 679)			(3 278; 22 337)	
WMT	5 264 919	1 614 599	6 433	12 456 917	3 290 958	13 111	26 919 968	4 266 108	16 862	22 534 757	3 449 871	13 690	11 566 208	1 888 542	10 045
		(3 037; 10 200)			(4 236; 27 298)			(4 728; 39 818)			(2 897; 44 283)			(3 907; 24 949)	
XOM	7 873 641	2 946 807	11 740	20 516 820	6 059 941	24 143	42 366 522	6 983 246	27 602	33 762 439	5 306 792	21 059	22 049 321	3 068 502	16 322
		(5 738; 35 560)			(6 030; 51 552)			(10 220; 70 619)			(5 818; 54 789)			(7 805; 57 633)	

Note: The table reports the yearly number of trades before and after cleaning and the average daily activity together with the minimum and maximum of number of trades between parentheses. The sample period is January 3 2006 to September 30 2010.

AA- Alcoa; AXP- American Express; BA- Boeing Company; CAT- Caterpillar; GE- General Electric; HD- Home Depot; IBM- Int. Business Machines; JPM- JPMorgan Chase & Co; KO- Coca-Cola; MCD- McDonald's; PFE- Pfizer; PG- Procter & Gamble; T- AT&T; WMT- Wal-Mart Stores; XOM- Exxon Mobil.

average number of clean transactions for all equities per day amounts to around 12 000.

Intraday returns can be constructed using different types of sampling schemes, see Zhou (1992), Hansen and Lunde (2006), Oomen (2006) and Griffin and Oomen (2008). Most common is the *calendar time sampling* scheme, where data is sampled at fixed time intervals. This scheme has a shortcoming that information can be insufficient in highly volatile intraday time intervals and redundant at other moments of the day. Alternatively, one can apply a *transaction time sampling* scheme, such that the observations are sampled every kth transaction. This sampling scheme adapts naturally to the trading activity, and thus estimates have lower variability relative to estimates based on data sampled in calendar time; see Oomen (2006). In our application however, we use a sampling scheme that forms a series of de-volatilized intraday returns, through a sampling Oomen (2006) refers to as *business time sampling*. This scheme dates back to Zhou (1992) who proposes to sample data each time some pre-specified amount of variation has realized. The resulting series is therefore equally informative. This scheme requires an estimate of intraday cumulative volatility. Since volatility is latent, we follow the idea of Fukasawa (2010) by applying cumulative volume as a proxy for cumulative latent volatility.[3] Thus, our sampling times $t_{0,N}, t_{1,N}, \ldots, t_{N,N}$ are such that amount of realized volume in each interval $(t_{i-1,N}, t_{i,N}]$ equals $(\textit{total volume})/N$, and we set $N \in \{250, 500\}$. On average, this corresponds to sampling every 48th and 24th transaction. Relative to calendar time sampling, we sample more frequently at the start and at the end of the trading day, while less frequently in the middle of the day. Our sample sizes also yield a reasonable compromise between potential microstructure effects and problems due to too small sample size. Actually, since the data is sampled sparsely, microstructure noise and price discreteness (as judged by first order serial correlation, volatility signature plots and histogram of price increments) are not expected to affect our results greatly.

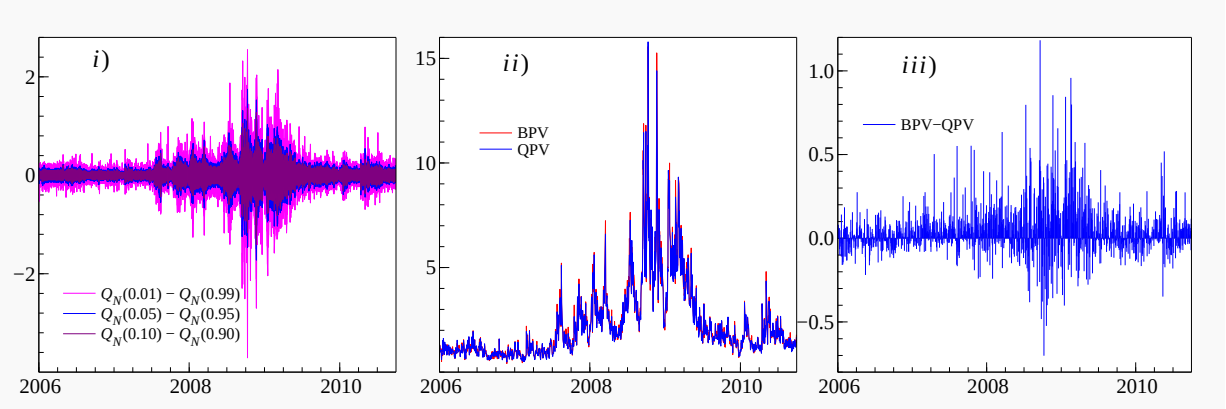
3.5.2 Empirical results

We start by illustrating the time series of computed empirical quantiles. Panel i) of Figure 3.5 plots symmetric pairs of quantiles of intraday returns for the JPM equity. We observe a considerably larger spread of empirical quantiles around 2008 and 2009, the period corresponding to the subprime crisis. Panels ii) and iii) plot the volatility estimates of σ_t resulting from the bipower and quantile-based variation measures for all days in series, and their difference. The two measures correspond closely, although a more detailed inspec-

[3]Alternatively, we also experimented using estimators of integrated variance robust to microstructure noise and jumps to obtain an estimate of cumulative variance. This delivered very similar results that do not change our conclusions. However, since the high-frequency based estimators depend on tuning parameters, and in general are considerably more complicated to obtain than cumulative volume, we report results based on the latter proxy only.

tion indicates that for several days the bipower variation deviates from the quantile-based measure. We observe that for these days the requirement that large price increments are preceded and succeeded by very small returns is not satisfied. This may in turn lead to loss of power of the BNS tests for these days.

Figure 3.5: *Empirical quantiles and variation measures*



We compute the ratio tests using intraday data from the trading days. As such, each day and each equity is treated on its own. Table 3.6 reports the fraction of days where one or more significant jumps are detected using the 5% significance level. The results can be summarized as follows. Using only $N = 250$ returns per day the rejections of the null hypothesis of no jumps are less frequent than for the sample size $N = 500$, as was also expected from the Monte Carlo results. At $N = 250$, the fraction of days where the second moment is affected by presence of jumps is around 20-40%, and it increases to 40-60% for the third and fourth moment as judged by the BJ test. At $N = 500$, we find that the fraction of days with the second moment significantly affected by jumps is roughly 40-60%, and it increases to 70-80% for the BJ$(\cdot, 3)$ and BJ$(\cdot, 4)$ tests. The sixth moment of the data seems to be less affected, although the Monte Carlo study suggests one should not fully trust results based on $r = 6$ at this sample size as the QPV may still suffer from a small bias due to the jumps, and the test statistic may be incorrectly sized. It is seen that for all three cases of $p \in \{1, 2, 3\}$ the results are similar. We do not report results when $p > 3$ sub-estimators are used in the quantile-based measure, as the optimal quantile q_1 may be too extreme in this case, cf. Table 3.1. Results not reported here indicate strong sensitivity with respect to the level of q_1 when $p > 3$, a result that is not satisfactory from the practical point of view.

The BJ$(\cdot, 2)$ and BNS tests provide in general similar results, which is in line with expectations from the simulations. Surprisingly, the JO test rejects the null less frequently than the BJ$(\cdot, 3)$ and BNS tests, which suggests that both positive and negative jumps

occur on the same day, offsetting each other. Indeed, a more detailed inspection of the data suggests that abnormally large price increments appear in both directions. We find that the fraction of days with identified jumps is slightly higher than reported by Gilder (2009) in his extensive study (based on the approach proposed by Andersen, Bollerslev, and Dobrev (2007)), also using US equities but over the period 2002-2006. As our sample period spans the financial crisis of 2007-2010, the jump process can be expected to be more active than in the earlier period as studied by Gilder (2009).

We study a few cases in more detail, in a similar way as Jiang and Oomen (2008). Figure 3.6 presents the scatter plot of the p-values of the test statistics considered for the JPM equity. Dotted lines present the nominal size of 5%. The lower left quadrant indicates instances where both the BJ and the adjusted BNS respectively the JO detect presence of jumps. The top left quadrant indicates when the BJ detects outliers while the alternative does not, while the bottom right quadrant shows the reverse. There is a substantial number of instances where the testing results overlap, but also when they lead to opposite decisions. It is of interest to have a closer look at these cases. Specifically, we look at the most extreme instances as measured by the distance between the p-values of the two competing tests, marked by circles in top-left and bottom-right corners of the panels of Figure 3.6.

Figure 3.6: *Scatter plots of p-values for the JPM equity*

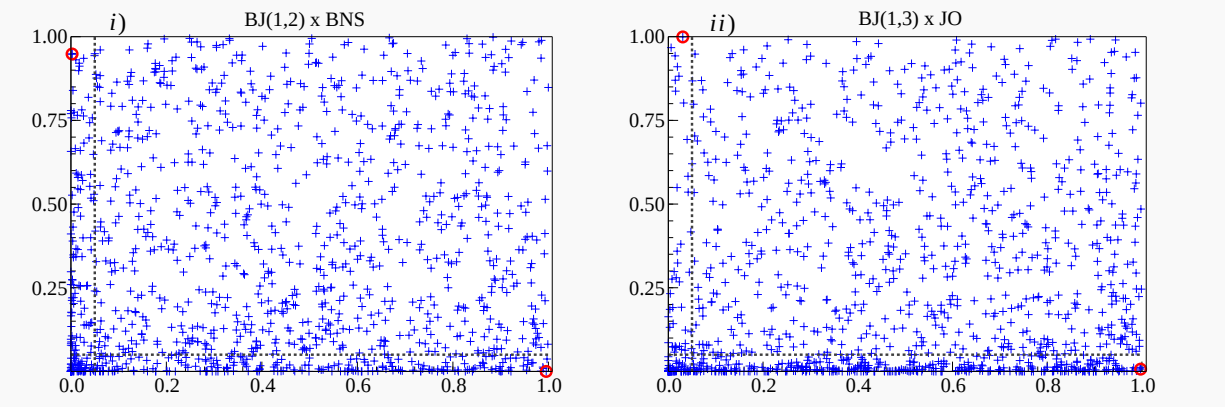


Table 3.7 reports p-values for the ratio tests for these most extreme instances. We focus on the second moment of the data and contrast BJ and BNS tests. Let us consider the January, 11, 2007 case. We observe that for $r = 2$ and $p \in \{1, 2\}$ the BJ ratio tests detect one or more outliers, while (adj)BNS but also JO do not. Only when $p = 3$ (and using the quantiles prescribed by the asymptotic minimum variance criterion used throughout the article), the BJ test does not detect significant jumps. If we lower the value of q_1 by e.g. 1%, a drop of the p-value is found (result not reported in the table). The top panel of

Table 3.6: *Testing results for equity data*

		$N = 250$								$N = 500$							
Symbol	BJ(p, ·)	BJ(·, 1)	BJ(·, 2)	BJ(·, 3)	BJ(·, 4)	BJ(·, 6)	adjBNS	BNS	JO	BJ(·, 1)	BJ(·, 2)	BJ(·, 3)	BJ(·, 4)	BJ(·, 6)	adjBNS	BNS	JO
AA	BJ(1,·)	0.3615	0.3699	0.5121	0.5322	0.1381	0.2126	0.2167	0.0904	0.6084	0.6025	0.7381	0.7808	0.6636	0.4502	0.4510	0.0962
	BJ(2,·)	0.5238	0.4008	0.5490	0.5481	0.2167	-	-	-	0.7356	0.6234	0.8000	0.8310	0.6619	-	-	-
	BJ(3,·)	0.6837	0.3623	0.5205	0.5222	0.1774	-	-	-	0.8167	0.5916	0.7958	0.8218	0.6184	-	-	-
AXP	BJ(1,·)	0.3556	0.3205	0.5029	0.5347	0.1824	0.1590	0.1632	0.1197	0.5423	0.5347	0.7749	0.8251	0.6987	0.3238	0.3238	0.1356
	BJ(2,·)	0.5715	0.3029	0.5397	0.5649	0.2418	-	-	-	0.7481	0.5531	0.8142	0.8577	0.7121	-	-	-
	BJ(3,·)	0.7356	0.3021	0.4987	0.5172	0.2134	-	-	-	0.8653	0.5029	0.8318	0.8619	0.6795	-	-	-
BA	BJ(1,·)	0.2996	0.3565	0.6008	0.6368	0.2753	0.1548	0.1565	0.1582	0.5063	0.5548	0.8293	0.8745	0.7816	0.3088	0.3088	0.1582
	BJ(2,·)	0.5657	0.3498	0.6000	0.6360	0.3331	-	-	-	0.7674	0.5289	0.8619	0.8979	0.7741	-	-	-
	BJ(3,·)	0.7431	0.2937	0.5724	0.6100	0.2795	-	-	-	0.8946	0.4828	0.8745	0.8996	0.7548	-	-	-
CAT	BJ(1,·)	0.3665	0.2527	0.4418	0.4678	0.1356	0.1573	0.1590	0.1096	0.5540	0.4176	0.7079	0.7598	0.6360	0.3088	0.3088	0.1046
	BJ(2,·)	0.5782	0.2301	0.4778	0.5038	0.1816	-	-	-	0.8234	0.4109	0.7297	0.7816	0.6259	-	-	-
	BJ(3,·)	0.7397	0.2176	0.4351	0.4661	0.1523	-	-	-	0.8921	0.3967	0.7782	0.8226	0.6092	-	-	-
GE	BJ(1,·)	0.3799	0.3883	0.4276	0.4067	0.1013	0.2485	0.2561	0.0762	0.6435	0.6485	0.6887	0.6703	0.5297	0.5272	0.5305	0.0720
	BJ(2,·)	0.4812	0.4134	0.4360	0.3950	0.1305	-	-	-	0.6736	0.7088	0.7506	0.7255	0.5381	-	-	-
	BJ(3,·)	0.5715	0.4561	0.4427	0.3858	0.1071	-	-	-	0.7255	0.6837	0.7238	0.7213	0.4862	-	-	-
HD	BJ(1,·)	0.2912	0.3791	0.5464	0.5682	0.2000	0.1623	0.1657	0.1013	0.5197	0.6159	0.7757	0.8201	0.7180	0.3238	0.3238	0.1138
	BJ(2,·)	0.4979	0.3531	0.5623	0.5866	0.2569	-	-	-	0.6837	0.6042	0.7933	0.8301	0.6845	-	-	-
	BJ(3,·)	0.6603	0.3297	0.5397	0.5531	0.2084	-	-	-	0.7808	0.5816	0.7933	0.8268	0.6711	-	-	-
IBM	BJ(1,·)	0.2485	0.2720	0.4519	0.4728	0.1406	0.1155	0.1197	0.1054	0.4603	0.4243	0.7029	0.7640	0.6134	0.2477	0.2485	0.1188
	BJ(2,·)	0.4510	0.2427	0.4594	0.4711	0.1967	-	-	-	0.7021	0.4368	0.7473	0.7799	0.6176	-	-	-
	BJ(3,·)	0.6393	0.1925	0.4159	0.4435	0.1515	-	-	-	0.8444	0.4042	0.7732	0.8176	0.5925	-	-	-
JPM	BJ(1,·)	0.2669	0.2393	0.3573	0.3615	0.0962	0.1381	0.1456	0.0962	0.4812	0.4159	0.5941	0.6544	0.5029	0.2527	0.2527	0.0962
	BJ(2,·)	0.4042	0.2820	0.3766	0.3774	0.1389	-	-	-	0.6351	0.4770	0.6611	0.6921	0.5230	-	-	-
	BJ(3,·)	0.5473	0.2485	0.3356	0.3331	0.1247	-	-	-	0.7456	0.4377	0.6611	0.7013	0.5004	-	-	-
KO	BJ(1,·)	0.3339	0.4100	0.5724	0.5808	0.2259	0.1515	0.1531	0.1038	0.5498	0.6059	0.7883	0.8234	0.7222	0.3021	0.3021	0.0962
	BJ(2,·)	0.5331	0.3975	0.5632	0.5682	0.2837	-	-	-	0.7113	0.6075	0.7941	0.8360	0.6996	-	-	-
	BJ(3,·)	0.6887	0.3448	0.5381	0.5515	0.2285	-	-	-	0.7908	0.6234	0.8226	0.8460	0.7146	-	-	-
MCD	BJ(1,·)	0.3481	0.3849	0.6017	0.6268	0.2192	0.2126	0.2142	0.1255	0.5506	0.5983	0.7824	0.8377	0.7431	0.3749	0.3749	0.1331
	BJ(2,·)	0.5749	0.3808	0.5967	0.6017	0.2904	-	-	-	0.7389	0.5908	0.8218	0.8544	0.7372	-	-	-
	BJ(3,·)	0.7556	0.3347	0.5339	0.5707	0.2393	-	-	-	0.8377	0.6017	0.8377	0.8753	0.6996	-	-	-
PFE	BJ(1,·)	0.4151	0.5105	0.5515	0.5439	0.1824	0.2787	0.2862	0.0987	0.6937	0.7406	0.7573	0.7448	0.6368	0.5967	0.5983	0.0828
	BJ(2,·)	0.5364	0.5297	0.5883	0.5464	0.2343	-	-	-	0.7004	0.7950	0.8050	0.7975	0.6628	-	-	-
	BJ(3,·)	0.5992	0.5724	0.5874	0.5540	0.1891	-	-	-	0.7565	0.7523	0.7849	0.7674	0.5799	-	-	-
PG	BJ(1,·)	0.3230	0.3615	0.5381	0.5582	0.2251	0.1841	0.1874	0.1339	0.5063	0.5858	0.7640	0.8100	0.7230	0.3063	0.3079	0.1255
	BJ(2,·)	0.4954	0.3682	0.5607	0.5799	0.2820	-	-	-	0.6862	0.6008	0.8075	0.8418	0.7063	-	-	-
	BJ(3,·)	0.6979	0.3448	0.5163	0.5448	0.2301	-	-	-	0.8276	0.5598	0.8134	0.8385	0.6862	-	-	-
T	BJ(1,·)	0.3439	0.4469	0.5791	0.5983	0.2318	0.2494	0.2552	0.1431	0.6084	0.6695	0.7900	0.8218	0.7339	0.4603	0.4611	0.1397
	BJ(2,·)	0.5071	0.4619	0.6050	0.6084	0.3038	-	-	-	0.6820	0.7063	0.8268	0.8435	0.7238	-	-	-
	BJ(3,·)	0.6343	0.4569	0.5632	0.5749	0.2494	-	-	-	0.7682	0.6912	0.8318	0.8527	0.7004	-	-	-
WMT	BJ(1,·)	0.2544	0.3347	0.4720	0.4812	0.1481	0.1255	0.1289	0.0862	0.4912	0.5289	0.7264	0.7808	0.6510	0.3029	0.3063	0.1079
	BJ(2,·)	0.4494	0.3372	0.4728	0.4703	0.2059	-	-	-	0.6460	0.5531	0.7607	0.7891	0.6502	-	-	-
	BJ(3,·)	0.5774	0.3163	0.4326	0.4444	0.1640	-	-	-	0.7657	0.5590	0.7640	0.8084	0.6234	-	-	-
XOM	BJ(1,·)	0.1983	0.1732	0.2435	0.2586	0.0703	0.0996	0.1038	0.0762	0.3833	0.2979	0.4611	0.5054	0.3732	0.2159	0.2184	0.1054
	BJ(2,·)	0.3029	0.1707	0.2494	0.2586	0.1013	-	-	-	0.5573	0.3565	0.5013	0.5272	0.3749	-	-	-
	BJ(3,·)	0.4167	0.1699	0.2410	0.2259	0.0954	-	-	-	0.6285	0.3682	0.5456	0.5682	0.3766	-	-	-

Figure 3.7a zooms in on this date, January 11, 2007. We observe that the largest negative price increment is neither preceded nor followed by very small returns. In such a situation, the bi- and quad-power variation can loose robustness, leading to a lower value of the BNS test statistic. This can explain the advantage of the BJ test over the BNS test.

Table 3.7: *Ratio tests for the selected case studies*

	BJ(·, 1)	BJ(·, 2)	BJ(·, 3)	BJ(·, 4)	BJ(·, 6)	BNS	adjBNS	JO
11Jan2007								
$p = 1$	0.0145	0.0000	0.0003	0.0052	0.1137	0.9954	0.9954	0.7951
$p = 2$	0.0001	0.0002	0.0871	0.4741	0.8320	-	-	-
$p = 3$	0.1019	0.5146	0.1453	0.2059	0.4908	-	-	-
17Jun2010								
$p = 1$	0.0658	0.9485	0.3765	0.3422	0.5771	0.0022	0.0026	0.1761
$p = 2$	0.0002	0.1138	0.9455	0.8094	0.9915	-	-	-
$p = 3$	0.0783	0.1304	0.0260	0.0925	0.4639	-	-	-
18Jun2009								
$p = 1$	0.0773	0.2277	0.0070	0.0050	0.0641	0.1480	0.1480	0.9973
$p = 2$	0.0280	0.0333	0.0001	0.0003	0.0375	-	-	-
$p = 3$	0.0000	0.6917	0.0009	0.0008	0.0436	-	-	-
08Dec2008								
$p = 1$	0.0244	0.2536	0.9994	0.4636	0.2620	0.8755	0.8816	0.0306
$p = 2$	0.4883	0.2280	0.0371	0.0306	0.1011	-	-	-
$p = 3$	0.1171	0.6097	0.0666	0.0404	0.1060	-	-	-

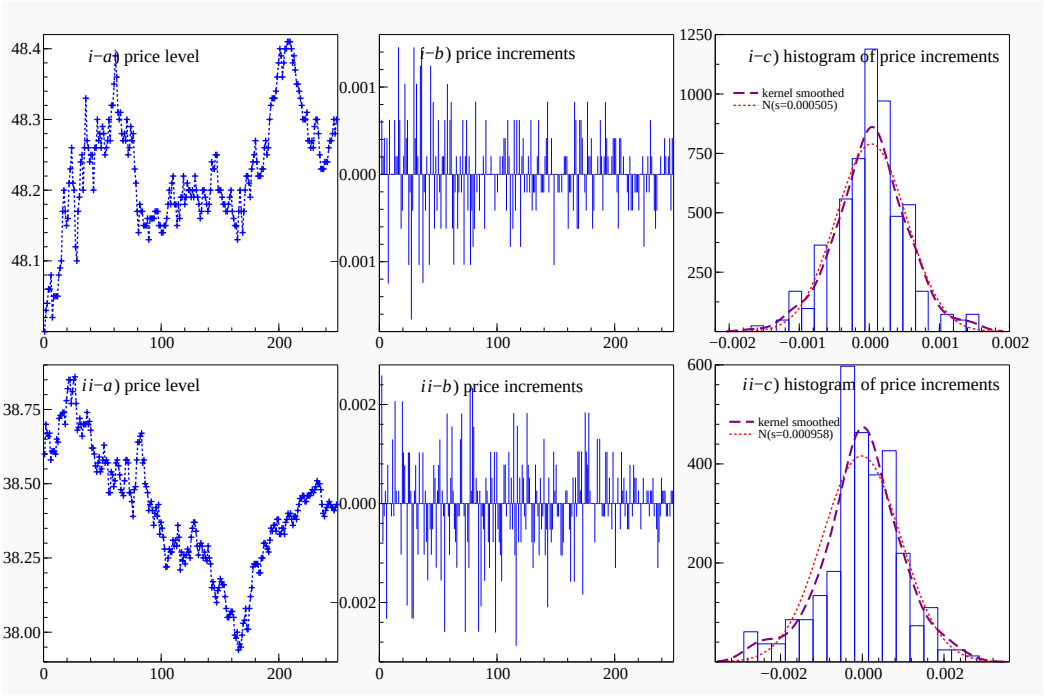
Note: The table reports p-values for the indicated tests.

For the June 17, 2010 case, the BJ ratio tests do not identify jumps for $r = 2$ and $p \in \{1, 2, 3\}$, while (adj)BNS does. Here the BNS and adjusted BNS test have different p-values which suggests that de-volatilization procedure was not entirely successful, and this may lead to worse performance of the BJ test. The bottom panel of Figure 3.7a focuses on this day, and we notice a cluster of larger negative price increments at the beginning of the sample, and a histogram where also the central part is distinctly non-normal. This is a case where the BJ test may fail.

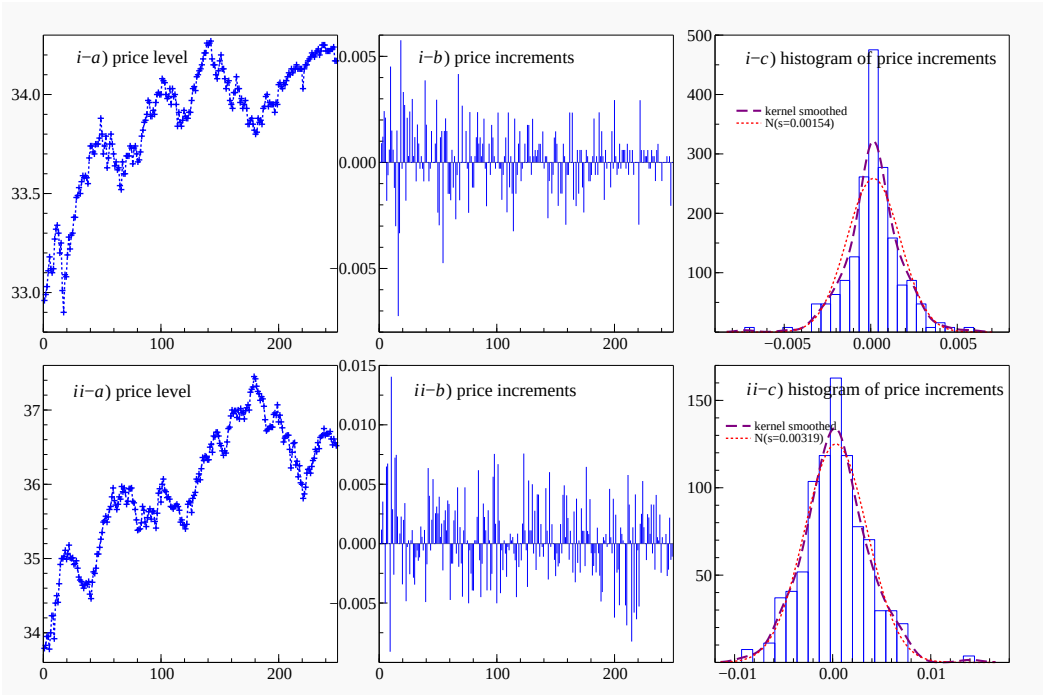
We now focus on the case when the test is built on the third moment and we contrast the BJ and JO tests. On June 18, 2009 the BJ ratio tests with $r = 3$ and $p \in \{1, 2, 3\}$ detect jumps, while the JO test does not. The top panel of Figure 3.7b zooms on June 18, 2009. We observe that a few positive as well as negative substantial price increments occur, particularly at the beginning of the sample. If there are multiple abnormal price increments of opposite signs, then the JO test can fail to detect them. It is the scenario where the BJ test has an advantage over the JO test.

On December 8, 2008 the JO test indicates the presence of jumps while the BJ$(1, 3)$ does not. The bottom panel of Figure 3.7b displays the returns for this date. There is a large positive price increment at the 15th observation, which also shows up in the right tail of the density as seen in panel *ii-c*) of Figure 3.7b. On this particular day, the empirical density in panel *ii-c*) displays relatively short tails (apart from the one outlier), which may have led to a downward bias on the QPV estimator. However, as soon as we combine more pairs, the BJ test lead to similar testing outcome as the alternative JO test.

Figure 3.7: Case studies for the JPM equity.



(a) The figure displays price levels, returns, and histogram/density of returns; top panel 11Jan2007, bottom panel 17Jun2010.



(b) The figure displays price levels, returns, and histogram/density of returns; top panel 18Jun2009, bottom panel 08Dec2008.

3.6 Summary and Conclusions

This chapter introduces the quantile-based measure of power variation of returns, which is robust against outlying (or abnormal) observations. We derive the joint asymptotic distribution of the quantile and moment-based measures of σ^r under the normal distribution. With the limiting distribution, we provide a new class of test statistics designed to detect outliers that lead to deviations from normality. The test statistics exploit different powers r, and hence facilitate judging the impact of outliers on different moments of the data. The tests are easy to implement and can be applied to any homoskedastic dataset, where it is conjectured that the presence of a finite number of outliers leads to non-normality.

We apply our theory to test for the presence of jumps in equity data over a sequence of trading days. In order to do so, we use the cumulative volume of the trades as a proxy for cumulative volatility, allowing us to devolatilize the returns to a homoskedastic series. Throughout the chapter, we compare the performance of our tests to widely applied alternatives: the tests of Barndorff-Nielsen and Shephard (2006) and of Jiang and Oomen (2008). The simulation results show that for higher orders of r our tests are more powerful than the alternative testing approaches. The empirical illustration shows that jumps can occur more frequent than judged by second or third moment of the data only.

In the empirical study, we alleviate the presence of microstructure noise in the high-frequency data by sampling the data sparsely. On the one hand, the lower frequency decreases the ability to detect jumps. On the other hand, problems related to price discreteness and other microstructure effects can also adversely effect the quantile-based and moment-based estimators of σ^r, and these problems are evaded here by the sparse sampling. Application of the same techniques on ultra high-frequency data, possibly using noise-reduction techniques as the pre-averaging of price increments proposed by Jacod, Li, Mykland, Podolskij, and Vetter (2009b), requires an extension of the asymptotic theory presented here, to allow for the dependence in data introduced by the noise reduction techniques. We leave this option for future research.

3.A Appendix

3.A.1 Proofs

Proof of Proposition 1: The proof follows the results stated in Theorem 3.2.1 and Corollary 3.2.2. Here we explain the steps in more detail. Let y be a normal variable with mean μ and variance σ^2 and let z be a standard normal variable. We make use of the symmetry property of the normal distribution and the convenient standard form,

$$f(y;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{ -\frac{(y-\mu)^2}{2\sigma^2} \right\} = \frac{1}{\sigma}\phi(z), \tag{3A.1}$$

where ϕ denotes the *pdf* of the standard normal variable. The proof consists of two parts. First we present the asymptotic variance-covariance of sample quantiles, then we derive the expressions for the partial derivatives of the function of quantiles that constitutes the quantile-based estimator (3.16).

We start by considering the $2p \times 1$ vector of quantiles $\vec{\boldsymbol{q}} = [\vec{q}_1, \ldots, \vec{q}_{2p}]' = [1-q_1, 1-q_2, \ldots, 1-q_p, q_p, \ldots, q_2, q_1]'$ with $1 > q_1 > \ldots > q_p > 1/2$. Using Theorem 3.2.1 and the symmetry of the density (3A.1), it follows that

$$N^{1/2}\left(Q_N(\vec{\boldsymbol{q}}) - Q(\vec{\boldsymbol{q}})\right) \xrightarrow{d} \mathcal{N}\Big(0, \sigma^2 \boldsymbol{G}\Big),$$

where $\boldsymbol{G}$ is the $2p \times 2p$ variance-covariance matrix given in (3.17).

We represent the $QPV_N^r(\boldsymbol{q}, \boldsymbol{\lambda}, p)$ estimator as the function $\varphi(\cdot,\cdot)$ of the vector $Q_N(\boldsymbol{q})$ and a given weighting scheme $\boldsymbol{\lambda}$. In particular,

$$\begin{aligned} QPV_N^r(\boldsymbol{q}, \boldsymbol{\lambda}, p) &= \varphi(Q_N(\boldsymbol{q}), \boldsymbol{\lambda}) \\ &= \boldsymbol{\lambda}' \left(\left[QPV_N^r(q_1), \quad \cdots, \quad QPV_N^r(q_p) \right]' \right), \end{aligned}$$

with $QPV_N^r(q)$ defined in (3.14). Following Corollary 3.2.2 we take the partial derivatives of function $\varphi(Q_N(\boldsymbol{q}), \boldsymbol{\lambda})$ with respect to $\vec{Q}_N = [Q_N(1-q_1), Q_N(1-q_2), \ldots, Q_N(q_2), Q_N(q_1)]$ and obtain

$$\frac{\partial QPV_N^r(\boldsymbol{q}, \boldsymbol{\lambda}, p)}{\partial \vec{Q}_N} = \boldsymbol{\lambda}' \begin{bmatrix} \frac{\partial QPV_N^r(q_1)}{\partial Q_N(1-q_1)} & \frac{\partial QPV_N^r(q_1)}{\partial Q_N(1-q_2)} & \cdots & \frac{\partial QPV_N^r(q_1)}{\partial Q_N(q_2)} & \frac{\partial QPV_N^r(q_1)}{\partial Q_N(q_1)} \\ \frac{\partial QPV_N^r(q_2)}{\partial Q_N(1-q_1)} & \frac{\partial QPV_N^r(q_2)}{\partial Q_N(1-q_2)} & \cdots & \frac{\partial QPV_N^r(q_2)}{\partial Q_N(q_2)} & \frac{\partial QPV_N^r(q_2)}{\partial Q_N(q_1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial QPV_N^r(q_p)}{\partial Q_N(1-q_1)} & \frac{\partial QPV_N^r(q_p)}{\partial Q_N(1-q_2)} & \cdots & \frac{\partial QPV_N^r(q_p)}{\partial Q_N(q_2)} & \frac{\partial QPV_N^r(q_p)}{\partial Q_N(q_1)} \end{bmatrix}'. \tag{3A.2}$$

Note that

$$\frac{\partial QPV_N^r(q_k)}{\partial Q_N(1-q_{k'})} = \frac{\partial QPV_N^r(q_k)}{\partial Q_N(q_{k'})} = 0 \qquad \text{for} \quad k \neq k', \tag{3A.3}$$

hence the matrix of derivatives in (3A.2) simplifies to a diagonal matrix.

We also obtain

$$\frac{\partial QPV_N^r(q_k)}{\partial Q_N(1-q_k)} = \frac{1}{c^r(q_k)} \sum_{j=0}^{r} (-1)^j j \binom{r}{j} \big[Q_N(q_k)\big]^{r-j} \big[Q_N(1-q_k)\big]^{j-1} \mathbf{1}_{\{j-1\geq 0\}},$$

$$\frac{\partial QPV_N^r(q_k)}{\partial Q_N(q_k)} = \frac{1}{c^r(q_k)} \sum_{j=0}^{r} (-1)^j (r-j) \binom{r}{j} \big[Q_N(q_k)\big]^{r-j-1} \big[Q_N(1-q_k)\big]^{j} \mathbf{1}_{\{r-j-1\geq 0\}},$$

where function $\mathbf{1}_{\{B\}}$ takes value one if condition B is true, and zero else. In the next step, we use the symmetry property $Q(q) = -Q(1-q)$ to arrive at

$$\frac{\partial QPV_N^r(q_k)}{\partial Q_N(1-q_k)} = -\frac{1}{c^r(q_k)} \big[Q(q_k)\big]^{r-1} \sum_{j=0}^{r} j \binom{r}{j}, \tag{3A.4}$$

$$\frac{\partial QPV_N^r(q_k)}{\partial Q_N(q_k)} = \frac{1}{c^r(q_k)} \big[Q(q_k)\big]^{r-1} \sum_{j=0}^{r} (r-j) \binom{r}{j} \equiv -\frac{\partial QPV_N^r(q_k)}{\partial Q_N(1-q_k)} \tag{3A.5}$$

as $\sum_{j=0}^{r} j\binom{r}{j} = \sum_{j=0}^{r} (r-j)\binom{r}{j}$.

The structure of (3.18) now follows from (3A.3), (3A.4) and (3A.5). This completes the proof. □

Proof of Proposition 2: In this proposition, it is the asymptotic covariance between the sample quantile and sample absolute moment (or its power) that is given. This asymptotic covariance term directly follows from the following theorem:

Theorem 3.A.1. *Let $Y = \{Y_i\}_{i=1}^N$ be a series of iid random normal variables with mean μ, standard deviation σ, probability distribution function f and inverse cumulative distribution function F^{-1}. Let $Q_N(q)$ denote the sample quantile of Y and let M_N denote the estimator of the sample mean of Y, i.e. $\mu_N = N^{-1}\sum_{i=1}^N Y_i$. The joint asymptotic distribution of $Q_N(q)$ and μ_N is given by the bivariate normal distribution*

$$N^{1/2}\big((Q_N(q), \mu_N)' - (Q(q), \mu)'\big) \xrightarrow{d} \mathcal{N}\big(0,\, \sigma^2\, \boldsymbol{c}(q)\big),$$

where

$$\boldsymbol{c}(q) = \begin{bmatrix} \frac{q(1-q)}{[\phi(\Phi^{-1}(q))]^2} & \bar{M}_z \frac{q}{\phi(\Phi^{-1}(q))} - \frac{1}{\phi(\Phi^{-1}(q))} \int_{-\infty}^{\Phi^{-1}(q)} z\phi(z)\,dz \\ \bar{M}_z \frac{q}{\phi(\Phi^{-1}(q))} - \frac{1}{\phi(\Phi^{-1}(q))} \int_{-\infty}^{\Phi^{-1}(q)} z\phi(z)\,dz & \bar{M}_z^2 - (\bar{M}_z)^2 \end{bmatrix},$$

with $\bar{M}_z = \int_{-\infty}^{\infty} z\phi(z)dz$ with z being iid standard normal variable. Proof. See Lin, Wu, and Ahmad (1980). □

Lin et al. (1980) actually present a more general theorem for any distribution that is absolutely continuous with density f satisfying $f(F^{-1}(q) > 0$, mean μ, and finite variance σ^2. The entries to the covariance matrix Ω are location invariant. This basic asymptotic result extends to the vector of multiple quantiles we use, and to any (absolute) moment, which completes the proof. □

3.A.2 Alternative jump test statistics

This part reviews the tests of Barndorff-Nielsen and Shephard (2006) and Jiang and Oomen (2008). The presented test statistics are derived under the null of no jump in the price process. We first fix some notation. Let $\{P_i\}_{i=0}^N$ denote the asset prices and let $Y_i = \ln P_i$ denote the log price. Define $R_i = P_i/P_{i-1} - 1$ and $r_i = Y_i - Y_{i-1}$ as the simple and log return respectively. Then,

JO tests:
i) the linear test:

$$\frac{N\Big(SwV_N - RV_N\Big)}{\sqrt{\Omega_N^{SwV}(p)}} \xrightarrow{d} \mathcal{N}\big(0,\,1\big), \tag{3A.6}$$

ii) the logarithmic test:

$$\frac{N \cdot BPV_N\Big(\ln SwV_N - \ln RV_N\Big)}{\sqrt{\Omega_N^{SwV}(p)}} \xrightarrow{d} \mathcal{N}\big(0,\,1\big), \tag{3A.7}$$

iii) the ratio test:

$$\frac{N \cdot BPV_N\Big(1 - \big[RV_N/SwV_N\big]\Big)}{\sqrt{\Omega_N^{SwV}(p)}} \xrightarrow{d} \mathcal{N}\big(0,\,1\big), \tag{3A.8}$$

where

$$SwV_N = 2\sum_{i=1}^{N}(R_i - r_i), \tag{3A.9}$$

$$RV_N = \sum_{i=1}^{N} r_i^2, \tag{3A.10}$$

$$BPV_N = \mu_1^{-2}\frac{N}{N-1}\sum_{i=1}^{N-1}|r_i||r_{i+1}|, \tag{3A.11}$$

$$\Omega_N^{SwV}(p) = \frac{\mu_6}{9}\frac{N^3\mu_{6/p}^{-p}}{N-p+1}\sum_{i=0}^{N-p}\prod_{k=1}^{p}|r_{i+k}|^{6/4}, \tag{3A.12}$$

with $\mu_k = \mathbf{E}[|z|^k]$ with $z \sim \mathcal{N}(0,1)$ and $p = 4$ or $p = 6$.

BNS tests:
i) the linear test:

$$\frac{N^{1/2}\Big(BPV_N - RV_N\Big)}{\sqrt{\vartheta \cdot QV_N}} \xrightarrow{d} \mathcal{N}\big(0,\,1\big), \tag{3A.13}$$

ii) the logarithmic test:

$$\frac{N^{1/2} \cdot BPV_N\Big(\ln BPV_N - \ln RV_N\Big)}{\sqrt{\vartheta \cdot QV_N}} \xrightarrow{d} \mathcal{N}\big(0,\,1\big), \tag{3A.14}$$

iii-a) the ratio test:

$$\frac{N^{1/2} \cdot BPV_N\Big(\big[BPV_N/RV_N\big] - 1\Big)}{\sqrt{\vartheta \cdot QV_N}} \xrightarrow{d} \mathcal{N}\big(0,\,1\big), \tag{3A.15}$$

iii-b) the adjusted ratio test:

$$\frac{N^{1/2}\Big(\big[BPV_N/RV_N\big] - 1\Big)}{\sqrt{\vartheta \cdot \max\big(1, QV_N/(BPV_N)^2\big)}} \xrightarrow{d} \mathcal{N}\big(0,\,1\big), \tag{3A.16}$$

where $\vartheta = (\pi^2/4) + \pi - 5$ and where

$$QV_N = \mu_1^{-4} N \sum_{i=0}^{N-4} \prod_{k=1}^{4} |r_{i+k}|. \tag{3A.17}$$

The adjusted ratio test adapts for homoskedastic variance.

Chapter 4

Heavy-tailed density models with long memory dynamics for volatility and dependence

Abstract: This chapter develops a novel framework for the simultaneous time series modeling of financial volatility and dependence based on long memory dynamics and heavy-tailed densities. The dynamic specification of our model accounts for the properties of the multivariate series and is robust to outliers or jumps in the data. Our empirical study for equity data indicates that the degree of memory in volatilities is similar across equities and the degree of memory in correlations between returns varies significantly. The forecasts from our modeling framework are compared with high-frequency realized measures. We outperform well-known benchmark models convincingly.

4.1 Introduction

Different econometric approaches for analyzing time-varying covariances have recently been developed. The most popular approaches are multivariate generalizations of univariate volatility models. An overview of the multivariate generalized autoregressive conditional heteroskedasticity (MGARCH) model is given by Bauwens, Laurent, and Rombouts (2006) and Silvennoinen and Teräsvirta (2009). A review of the multivariate stochastic volatility (MSV) model is included in Asai, McAleer, and Yu (2006). Both model formulations are for multiple time series of, typically, daily returns. The MGARCH model belongs to the class of observation-driven models in which we formulate conditional densities. The MSV model belongs to the class of parameter-driven models in which we formulate unconditional joint densities.

Andersen, Bollerslev, Diebold, and Labys (2001) and Barndorff-Nielsen and Shephard (2002) discuss methods for computing realized measures for volatility, or the second moment of a variable, and for covariance, or the second moment across different variables. Realized measures are computed from high-frequency intra-daily returns. Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009) discuss kernel-based or nonparametric methods with different weighting schemes for the efficient estimation of the second moments using high-frequency data.

Andersen et al. (2001) argue further that realized measures behave as slowly mean-reverting or fractionally integrated processes. This empirical finding can justify the use of univariate fractionally integrated models for volatility as considered by Robinson (1991), Baillie, Bollerslev, and Mikkelsen (1996), Bollerslev and Mikkelsen (1996) and Tse (1998). The empirical evidence of long memory features in financial volatilities can have important implications for simultaneous option pricing as discussed in Rombouts and Stentoft (2010).

A number of studies have appeared where the dynamic processes in MGARCH models are designed to allow for long memory properties. Teyssière (1997) treats a bivariate fractionally integrated MGARCH for two exchange rate return series where the two volatilities and the covariance have similar long memory properties. A similar model is considered by Brunetti and Gilbert (2000) who analyze two crude oil price series. In the models of Teyssière (1997) and Brunetti and Gilbert (2000), the covariances are time-varying only due to the volatility series. The fractional integration orders for the two volatility processes are estimated as approximately equal values. Pafka and Mátyás (2001) extend Teyssière (1997) by considering a trivariate MGARCH model where the fractional integration coefficient is set equal for all six series associated with the 3×3 covariance matrix. Conrad, Karanasos, and Zeng (2010) apply bivariate and trivariate fractionally integrated MGARCH models with constant correlations to return series from eight major stock indexes and also they find that the fractional coefficients for the volatilities are estimated by approximately equal values.

The fractionally integrated MGARCH models used in the earlier studies have several drawbacks. Firstly, the time-varying conditional covariance matrix cannot be made subject to a set of analytical restrictions to preserve its positive definite property, see Teyssière (1997). At each time-update of the MGARCH model, we need to validate this property and amend the matrix when necessary. Therefore, the constant correlation assumption is often imposed. This assumption however is too restrictive in view of empirical studies; see Engle (2002a). Secondly, the model formulations are directed towards volatility. Parameter estimation is then subject to a set of nonnegativity constraints which may lead to some difficulties in estimation. Thirdly, Conrad et al. (2010) stress the necessity of considering heavy-tailed densities but for each series they adopt the same value for the heavy-tail coefficient. From the univariate analyses by Bollerslev (1987) we learn that tail properties can be very different for different financial series. In this chapter we aim to

alleviate these drawbacks and we therefore aim to develop a more suitable model for the simultaneous modeling of asset returns. The model accounts for conditional densities with heavy-tails and for long memory features in the time-varying volatilities and correlations. The dependence between the series is based on copula functions where marginal densities can be individually specified with different tail as well as memory coefficients.

It is argued in Joe and Xu (1996) that the copula approach provides a flexible modeling framework for multivariate data. Observation driven time-varying copulas have been originally developed by Patton (2002, 2006) while further work has been carried out by Dias and Embrechts (2004), Jondeau and Rockinger (2006), Bartram et al. (2007), Ausin and Lopes (2010) and Dias and Embrechts (2010). Hafner and Manner (2010) consider parameter driven counterparts of the dynamic copulas. In most of these studies, the time-varying dependencies have been estimated as highly persistent processes, see, in particular, Jondeau and Rockinger (2006, Table 6). Hafner and Manner (2010) even report that some copulas' dependence processes are close to a random walk.

We introduce long memory processes for volatility and dependence in a copula framework. The marginal volatility processes are modelled by univariate t densities and the dependence is modelled by a bivariate t copula density. The resulting conditional model is typically designed for multiple time series with heavy-tailed properties. For example, the choice of the t copula density is motivated by its ability to capture dependence of extreme events. Since the dynamic features of our modeling framework rely on the score of the t density, we are able to downweight possible outliers in the observation density as well as in the updating of the parameters over time. This robust property is crucial for long memory models where the impact of innovations on future values dissipates at a slow hyperbolic rate.

In our Monte Carlo study we show that our robust long memory specification for the dynamic copula outperforms short memory alternatives when the copula time dependence has a slow rate of decay. Our empirical study is based on a set of equity data. The high-frequency tick data is from the Trade and Quote (TAQ) database for the period January 1993 to May 2010 downloaded through the Wharton Research Data Services (WRDS) system. We construct daily log returns and realized volatility measures from the database. The realized measures are used to evaluate the quality of estimation across different models. We find empirical evidence of long memory in our estimated volatility series. The degree of long memory is similar to what is found in earlier studies. Long memory features also appear to be important for our dynamic t copula. The equity correlations rely significantly on fractionally integrated processes. The degrees of long memory for the correlations across pairs is substantially different and different from the degrees found in volatilities.

The remainder of the chapter is organized as follows. In Section 2 we introduce our long memory models for volatility and dependence. In Section 3 we conduct a Monte Carlo study to examine the performance of some competing models for modeling dependence un-

der a variety of correlation dynamics. In Section 4 we present the results of our empirical study. We relate model-based volatilities and correlations to high-frequency based realized measures and we discuss important differences of our models relative to competing alternatives. Finally, Section 5 concludes.

4.2 Conditional Volatility and Dependence

We provide a new dynamic model formulation for volatility in return series and for dependence between return series. Both volatility and dependence evolve over time as conditional dynamic processes that have long memory properties. The models rely on heavy-tailed densities. In the case of volatility we consider the Student's t density and for dependence we consider the bivariate t copula function.

4.2.1 Modeling of returns and volatility

Let y_t^* denote a daily log return series for time $t = 1, \ldots, n$. The expected return is denoted by θ and a consistent estimate of θ is the sample mean of y_t^* denoted by $\bar{y}^*$. Define $y_t = y_t^* - \bar{y}^*$ for $t = 1, \ldots, n$. The mean-adjusted log return y_t is modelled as a white noise series and each realization is assumed to come from a Student's t distribution with density given by

$$p(y_t|\sigma^2, \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{(\nu-2)\pi\sigma^2}}\left[1 + \frac{y_t^2}{(\nu-2)\sigma^2}\right]^{-(\nu+1)/2}, \tag{4.1}$$

with variance $\sigma^2 > 0$, number of degrees of freedom $\nu > 2$ and Gamma function $\Gamma()$. It is well known that the Gaussian distribution is obtained as a special case of the t distribution when $\nu \to \infty$ or $\nu^{-1} \to 0$.

It is common practice to model the variance of financial returns as time-varying functions in the form of generalized autoregressive conditional heteroskedasticity (GARCH) models of Engle (1982) and Bollerslev (1986). Another class of models that enables a time-varying variance is the stochastic volatility (SV) model of Taylor (1986) and Harvey, Ruiz, and Shephard (1994). We consider a scheme for modeling volatility that is related to both GARCH and SV models. It is however directly based on the Student's t distribution with density (4.1). The variance σ^2 is allowed to be time-varying and is denoted by σ_t^2 for $t = 1, \ldots, n$. To enforce the volatility process to have only positive support we rather model log volatility defined as $h_t = \log \sigma_t^2$.

The evolution of h_t over time is typically stationary and we can, for example, consider the first-order autoregressive process

$$h_{t+1} = \delta + \beta h_t + \alpha\eta_t, \qquad t = 1, \ldots, n, \tag{4.2}$$

where δ is an unknown constant, β is an autoregressive coefficient, α is a scaling constant and η_t is an innovation term with mean zero. For a stationary process, the unconditional mean of log variance h_t is given by $\mu = \delta/(1-\beta)$. The dynamics can be expressed in terms of mean-deviations, that is

$$h_{t+1} - \mu = \beta(h_t - \mu) + \alpha\eta_t, \qquad t = 1, \ldots, n. \tag{4.3}$$

The nature of the innovation term determines the class of model. For example, in case η_t is a stochastic disturbance term, the model is part of the SV class. In case we let η_t be a function of the current observation, say $\eta_t = y_t/\sigma_t$ we obtain a model that is related to the GARCH model. In particular, the model corresponds closely to the exponential GARCH (EGARCH) of Nelson (1992). In our framework we follow Creal, Koopman, and Lucas (2008) and let η_t to be a different function of the current observation. The details are given below.

It is also argued that the variance σ_t^2 or the log variance h_t typically follow high-persistence stationary processes; see Robinson (1991), Baillie et al. (1996) and Tse (1998). Most contributions propose long memory processes as means to model high-persistence volatility processes. In particular, the fractionally integrated process is often considered and the long memory formulation can be written as

$$(1-L)^d(h_{t+1} - \mu) = \beta(1-L)^d(h_t - \mu) + \alpha\eta_t \qquad t = 1, \ldots, n, \tag{4.4}$$

where L is the lag operator (with $Lh_t = h_{t-1}$) and the fractional difference operator $(1-L)^d$ can be defined by the binomial expansion

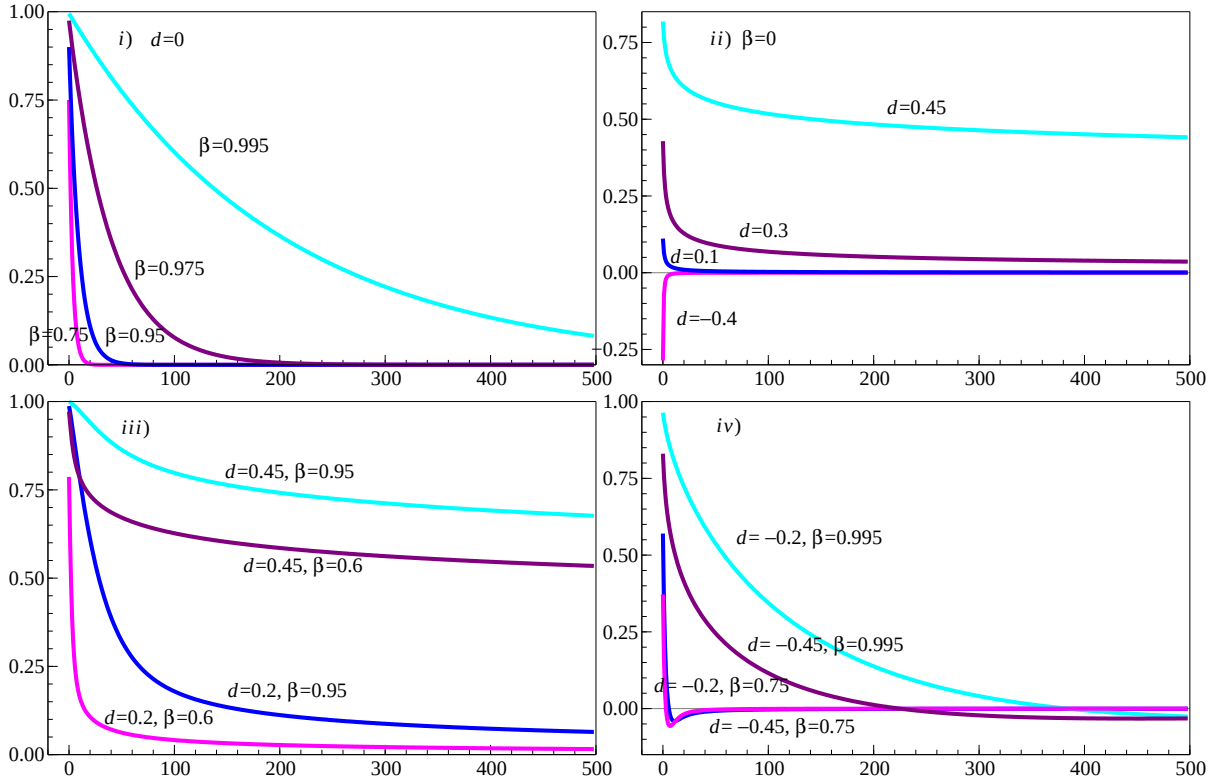
$$(1-L)^d = 1 - dL + \frac{d(d-1)}{2!}L^2 - \frac{d(d-1)(d-2)}{3!}L^3 + \ldots, \tag{4.5}$$

for any real order of fractional integration $d > -1$. To enforce a stationary process for h_t, one can impose $d < 1/2$. In empirical studies on volatility in financial markets, positive values are usually obtained for d.

The dynamic process (4.4) is within the class of ARFIMA processes as introduced by Granger and Joyeaux (1980) and Hosking (1981) for the conditional mean. Most statistical properties for the ARFIMA process are applicable to (4.4). By analogy, if autoregressive coefficient $|\beta| < 1$, fractional integration coefficient $d < 1/2$ and η_t is a martingale difference sequence with finite and time-invariant variance, a covariance stationary and invertible solution exists for (4.4); see Palma (2007, Section 3.2). The long memory process adds a considerably larger range of autocovariance functions compared to the short memory process (4.2); see Appendix 4.A.1 for its general structure.

Figure 4.1 present the autocorrelation function of long memory process. The memory or temporal dependence of the h_t process depends on β and d; compare panels i) to iv). For

Figure 4.1: *Autocorrelation functions for long memory model*



$d = 0$ autocovariance function decays at the exponential rate and the process is said to have short memory. For $d < 0$ the process has intermediate memory and all autocovariances (except for lag 0) are negative and decay hyperbolically to zero. For $d < 1/2$ the process is persistent and has long memory with an autocovariance function characterized by a slow hyperbolic rate of decay: the innovation η_t has a long-run effect on future values of h_t. Finally, for $1/2 \leq d < 1$ the process is not covariance stationary but it is still mean-reverting; see Baillie (1996, p.22). For $d = 1$ the log variance is a unit process.

A possible generalization of our formulation (4.4) is the introduction of the leverage effect. It can be incorporated by introducing an indicator function which makes the volatility respond differently to positive and negative shocks; see Glosten, Jagannathan, and Runkle (1993). We will discuss such a generalization in our empirical study in Section 4.4.

4.2.2 Dynamic conditional modeling of volatility

To determine the volatility innovation η_t in (4.2) or (4.4), we follow Creal et al. (2008). When considering the variance $h = \sigma^2$ directly, they propose to let η_t in (4.2) be equivalent

to the scaled score function for (4.1) with respect to σ^2 at time t. The score function is defined as the first derivative of log density of (4.1) with respect to σ^2 and is evaluated at $\sigma^2 = \tilde{\sigma}_t^2$ where we assume that $\tilde{\sigma}_t^2$ is available as it is computed in the previous time period $t-1$. The score function is scaled by the inverse of its asymptotic variance which is typically the expectation of the information matrix. We evaluate the score and expected information at $\sigma^2 = \tilde{\sigma}_t^2$. The details of computing η_t in this way for $h = \log \sigma^2$ are given below. We may regard this choice of η_t as ad-hoc. However, in the case of a normal density function for y_t, which is obtained by having $\nu^{-1} \to 0$ in (4.1), this approach leads to an updating scheme for the variance that is the same as in the GARCH model. Creal et al. (2008) show that when considering other densities, other well-known conditional dynamic models in financial econometrics are obtained using their framework. Also they argue that the score at time t is a natural candidate to adapt the current value of the parameter for the next period. When parameters are estimated by the method of maximum likelihood via direct numerical optimization, the score vector is also used as a basis for estimation. Finally, the martingale difference property of the score as we define it below also makes it a convenient choice for the innovation η_t in (4.2) or (4.4).

When we opt for the dynamic modeling of the log variance $h = \log \sigma^2$, the log density of (4.1) in terms of h is given by

$$\log p(y_t|h;\nu) = c(\nu) - \frac{1}{2}h - \frac{\nu+1}{2}\log\left(1 + \frac{y_t^2}{(\nu-2)\exp h}\right), \tag{4.6}$$

where $c(\nu)$ is a function of ν only. Differentiating the log density with respect to h (see Appendix 4.A.2 for details), we define the score at time t as

$$\nabla_t = \frac{\partial \log p(y_t|h;\nu)}{\partial h}, \tag{4.7}$$

and is given by

$$\nabla_t = \frac{1}{2}\left[\omega_t y_t^2/\sigma^2 - 1\right], \tag{4.8}$$

where the weight ω_t is given by

$$\omega_t = \frac{\nu+1}{(\nu-2) + y_t^2/\sigma^2} \geq 0. \tag{4.9}$$

For fixed values y_t and h, the weight ω_t converges to one as $\nu^{-1} \to 0$. Hence the weight is unity in the Gaussian case. We further learn that a smaller weight is obtained when the return y_t is too large in absolute terms with respect to its volatility σ. The expectation taken with respect to observation density $p(y_t|h;\nu)$ of the score ∇_t is zero by construction, that is $\mathbb{E}(\nabla_t) = 0$. We can refer to ∇_t as a martingale difference. The Fisher information at time t in our framework is defined as

$$\mathcal{I}_t = \mathbb{E}_{t-1}\left[(\nabla_t)^2\right], \tag{4.10}$$

and is given by

$$\mathcal{I}_t = \frac{1}{2}\frac{\nu}{\nu+3}. \tag{4.11}$$

We can conclude that the asymptotic variance of the score is a fixed scaling constant. In the limiting Gaussian case as $\nu \to \infty$, we obtain $\mathcal{I}_t = \frac{1}{2}$.

With the score and expected information, we make one-step ahead prediction of log volatility. For a given value of h at time t, that is $\tilde{h}_t$, we obtain the next value of h at time $t+1$ via (4.2), that is

$$\tilde{h}_{t+1} = \delta + \beta\tilde{h}_t + \alpha\tilde{\eta}_t, , \qquad t = 1, \ldots, n, \tag{4.12}$$

where our choice of $\tilde{\eta}_t$ is given by

$$\tilde{\eta}_t = \tilde{\mathcal{I}}_t^{-1}\tilde{\nabla}_t, \qquad \tilde{\nabla}_t = \nabla_t|_{h=\tilde{h}_t}, \qquad \tilde{\mathcal{I}}_t = \mathcal{I}_t|_{h=\tilde{h}_t}, \tag{4.13}$$

where ∇_t and $\mathcal{I}_t$ are defined above and where we treat δ, β and α as unknown coefficients that need to be estimated. Since $\tilde{\eta}_t$ is scaled by α and $\mathcal{I}_t$ is a fixed constant, we can also take

$$\tilde{\eta}_t = \tilde{\nabla}_t = \frac{1}{2}\big(\tilde{\omega}_t y_t^2/\exp\tilde{h}_t - 1\big), \tag{4.14}$$

where $\tilde{\omega}_t = (\nu+1)\,/\,\big((\nu-2) + y_t^2/\exp\tilde{h}_t\big)$. For a starting value of $h = \tilde{h}_1$, we can apply (4.12) for $t = 1, \ldots, n$ to obtain an updated estimate for log variance h for each time period. The estimation of the other parameters ν, δ, β and α is discussed in Section 4.2.5.

The general modeling framework of introducing time-varying parameters in model density specification is developed by Creal et al. (2008). A similar development as presented above is also developed by Harvey and Chakravarty (2008) who refer to their short memory model as the Beta-t-EGARCH model. Here we consider the long memory specification (4.4) for h_t and obtain

$$(1-L)^d(\tilde{h}_{t+1} - \mu) = \beta(1-L)^d(\tilde{h}_t - \mu) + \alpha\tilde{\eta}_t, \qquad t = 1, \ldots, n, \tag{4.15}$$

where $\tilde{\eta}_t$ is defined as in (4.14). The fractional integration version for the updating of the log variance introduces an additional unknown parameter, the fractional difference coefficient d, which needs to be estimated as well. Attractive features of this framework are the Student's t distribution for the demeaned log return series y_t and the long memory log volatility h_t process. The dynamic process for h_t is a function of the unknown parameters but also of the squared return y_t^2 and the time-varying weight ω_t. The weight ω_t is a direct consequence of considering the Student's t distribution for y_t. For a finite value of ν, ω_t decreases with an increasing value of y_t^2; it downweights the impact of more extreme observations. The weight (4.9) is bounded at a maximum value of $(\nu+1)/(\nu-2)$. Also, the weight ω_t is higher when ν is smaller. This robustness property distinguishes our modeling

framework from existing fractional integrated models for volatility, in particular from the fractionally integrated GARCH and EGARCH models of Baillie et al. (1996) and Bollerslev and Mikkelsen (1996).

A volatility modeling framework that is robust to large positive or negative log returns is important when the log volatility is modelled as a long memory process. The decay of large shock in the innovation $\tilde{\eta}_t$ is at a slow hyperbolic speed rather than at a geometric rate. The role of ω_t as downweighting y_t^2 when it is too large is therefore crucial for long memory log volatility as in (4.4). We provide empirical illustrations in Section 4.4 to illustrate the robustness property of our approach relative to existing long memory volatility models.

4.2.3 Modeling of dependence

The log return series y_t can be standardized by their volatilities $\tilde{\sigma}_t = \exp(\tilde{h}_t/2)$ obtained from above. We defined the standardized returns as

$$z_t = \frac{y_t}{\tilde{\sigma}_t}, \qquad t = 1, \ldots, n.$$

When considering a set of k time series of daily log return series, we define z_{it} as z_t but now it is the standardized returns for the ith daily return series with $i = 1, \ldots, k$. Copula functions can be used to describe the dependence structure between the series in a flexible way. We define u_{it} as the probability integral

$$u_{it} = \int_{-\infty}^{z_{it}} p_{z_i}(x)\mathrm{d}x,$$

where p_{z_i} denotes the marginal density of z_i. It follows that u_{it} is uniformly distributed, that is $u_{it} \sim U(0,1)$, for $i = 1, \ldots, k$. A copula function is defined generally by

$$C(u_{1t}, u_{2t}, \ldots, u_{kt}) = F\left[F_1^{-1}(u_{1t}),\, F_2^{-1}(u_{2t}),\, \ldots,\, F_k^{-1}(u_{kt})\right], \tag{4.16}$$

where $F_i()$ is the continuous cumulative distribution function (*cdf*) of the ith marginal variable for which the inverse function $F_i^{-1}()$ is assumed to exist and where $F()$ is the joint *cdf* for all series $z_{1t}, \ldots, z_{kt}$. Multivariate analysis based on copula functions is due to Sklar's theorem which states that any k-dimensional joint distribution function can be decomposed into its k marginal distributions and a copula function that completely describes its dependence structure. In our treatment below we consider the bivariate t copula density because it has the ability to capture also the dependence between extreme events; see the discussion in Demarta and McNeil (2005).

The dependence structure between two marginal series can be described by the bivariate t copula density as given by

$$c^T(u_{1t}, u_{2t}|\rho;\kappa) = \frac{\gamma(\kappa)}{(1-\rho^2)^{1/2}} \frac{\left[1+\kappa^{-1}(1-\rho^2)^{-1}\left(x_{1t}^2 + x_{2t}^2 - 2\rho x_{1t}x_{2t}\right)\right]^{-(\kappa+2)/2}}{\left[(1+x_{1t}^2/\kappa)(1+x_{2t}^2/\kappa)\right]^{-(\kappa+1)/2}}, \tag{4.17}$$

where

$$\gamma(\kappa) = \frac{\Gamma(\frac{\kappa+2}{2})\Gamma(\frac{\kappa}{2})}{\Gamma^2(\frac{\kappa+1}{2})}, \qquad x_{it} = T_\kappa^{-1}(u_{it}), \qquad i = 1, 2,$$

with correlation coefficient ρ, with uniform variable $u_{it} \in (0, 1)$ as defined above and with $T_\kappa()$ as the cumulative distribution function of the univariate t distribution with $\kappa > 0$ degrees of freedom. The resulting joint distribution of the bivariate data is only elliptical if the number of degrees of freedom in the t copula density (4.17) and the number of degrees of freedom of the two marginal densities (4.1) are equal. In particular, we have $\kappa = \nu_1 = \nu_2$ where ν_i is the degree of freedom of the marginal Student's t density for the ith series z_{it}, for $i = 1, 2$. In this case the joint bivariate distribution corresponds to the bivariate t distribution. The Gaussian copula density is obtained as the limiting case of the t copula density for $\kappa \to \infty$. The Gaussian copula density is given by

$$c^G(u_{1t}, u_{2t}|\rho) = (1-\rho^2)^{-1/2}\exp\Big[-\frac{1}{2}\frac{1}{1-\rho^2}\big(x_{1t}^2 + x_{2t}^2 - 2\rho x_{1t}x_{2t}\big) + \frac{1}{2}\big(x_{1t}^2 + x_{2t}^2\big)\Big], \quad (4.18)$$

where $x_{it} = \Phi^{-1}(u_{it})$, $i = 1, 2$, and where $\Phi()$ denotes the cumulative distribution function of the univariate Gaussian distribution. The Gaussian copula has zero dependence in the tails of the distribution.

The correlation coefficient ρ determines the dependence between two variables. For this purpose, we define the variable g as

$$g = \log\frac{1+\rho}{1-\rho}, \quad (4.19)$$

so that $-1 < \rho < 1$ for any real value of g. The variable g is allowed to vary over time and we denote it as g_t. We consider the autoregressive process for g_t as given by

$$g_{t+1} = \delta + \beta g_t + \alpha\varepsilon_t, \qquad t = 1, \ldots, n, \quad (4.20)$$

where unknown constant δ, autoregressive coefficient β and scaling constant α are deliberately chosen to have the same symbols as in (4.2) for the log volatility h_t. The coefficients have the same roles and their estimation for the volatility process h_t and for the dependence process g_t will be carried out separately. The process for g_t in (4.20) is imposed to be stationary. The innovation term ε_t has mean zero. The dependence can also be modelled as a fractionally integrated process. We then have

$$(1-L)^d(g_{t+1} - \mu) = \beta(1-L)^d(g_t - \mu) + \alpha\varepsilon_t, \qquad t = 1, \ldots, n, \quad (4.21)$$

where $\mu = \delta/(1-\beta)$. The discussion in Section 4.2.1 also applies to the dynamic process for dependence.

4.2.4 Dynamic conditional modeling of dependence

We determine the dependence innovation ε_t in (4.20) or (4.21) as in Creal et al. (2008) and Section 4.2.2. For the transformed dependence variable g in (4.19), we define ∇_t and $\mathcal{I}_t$ similarly as in (4.7) and (4.10), respectively, but here for the log of the t copula density (4.17) with respect to g, that is

$$\nabla_t = \frac{\partial \log c^T(u_{1t}, u_{2t}|g;\kappa)}{\partial g}, \qquad \mathcal{I}_t = \mathbb{E}_{t-1}[(\nabla_t)^2], \tag{4.22}$$

where the density dependence of ρ is replaced by g through (4.19). We obtain the score (see Appendix 4.A.2 for details) as given by

$$\nabla_t = \frac{\dot\rho}{(1-\rho^2)^2}\Big[(1+\rho^2)\big(\pi_t x_{1t}x_{2t} - \rho\big) - \rho\big(\pi_t x_{1t}^2 + \pi_t x_{2t}^2 - 2\big)\Big], \qquad \dot\rho = \frac{\partial \rho}{\partial g}, \tag{4.23}$$

where

$$\pi_t = \frac{\kappa+2}{\kappa+m_t} \geq 0 \qquad \text{with} \qquad m_t = \frac{1}{1-\rho^2}(x_{1t}^2 + x_{2t}^2 - 2\rho x_{1t}x_{2t}) \geq 0. \tag{4.24}$$

Here, m_t is the Mahalanobis distance of the bivariate vector $x_t = (x_{1t}, x_{2t})'$ to zero with respect to its covariance matrix

$$R = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

that generally equals $x_t' R^{-1} x_t$. To complete the model specification, we provide the expression for expected information. Since the correlation coefficient ρ appears only in the t density part of the t copula, the score and Fisher information of the correlation parameter in the t copula density correspond to those from the bivariate t density. The expected information reads

$$\mathcal{I}_t = \frac{\dot\rho^2}{(1-\rho^2)^2}\left(1 + \rho^2 - \frac{2\rho^2}{\kappa+2}\right)\frac{\kappa+2}{\kappa+4}. \tag{4.25}$$

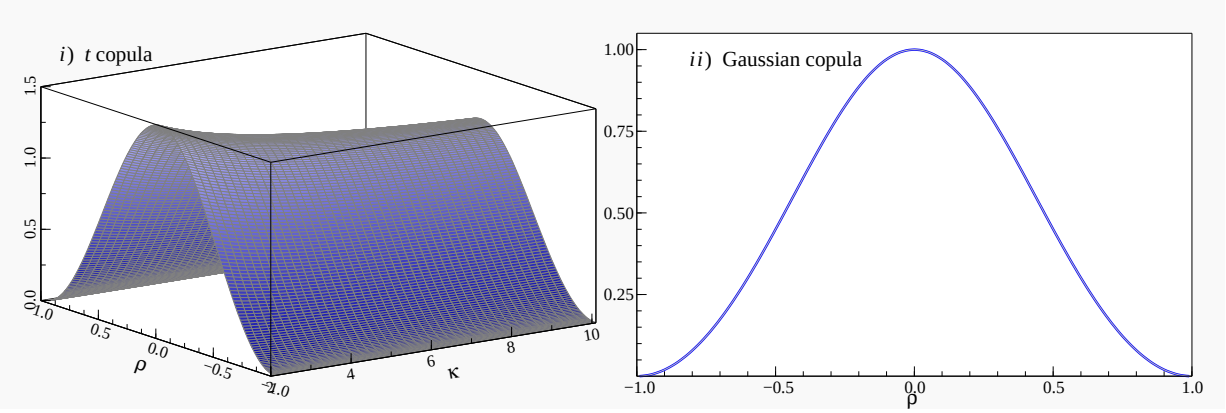
The variable π_t in (4.24) can be interpreted as a time-dependent weight assigned to the (product of) marginals. Outliers or t tail observations (either one or both) making the Mahalanobis distance m_t in (4.24) large are effectively downweighted for a finite value of degrees of freedom κ. The degree of robustness increases with decreasing κ that reflects tailness of the t copula density. The time-varying weight (4.24) is bounded from above at value $(\kappa+2)/\kappa$.

This updating scheme for the time-varying correlation based on the bivariate Gaussian copula density (4.18) yields the score ∇_t as in (4.23) but with $\pi_t = 1$ and it yields the Fisher information

$$\mathcal{I}_t = \dot\rho^2 \frac{1+\rho^2}{(1-\rho^2)^2}.$$

The model based on the bivariate Gaussian copula density is not robust against outliers since $\pi_t = 1$ implies that all observations, including large shocks, contribute to the update of the time-varying correlation.

Figure 4.2: *Inverse of Fisher information for t and Gaussian bivariate copulas*



Note: $\rho \in (-1, 1)$ and $\kappa \in (2, 10]$.

Figure 4.2 presents the shape of the inverse of the Fisher informations for t and Gaussian bivariate copula densities. The shape of inverse of the Fisher information depends on the (time-varying) correlation coefficient ρ. In the case of the t copula density it depends also on the number of degrees of freedom κ. It is seen that inverse of the Fisher information attains a small value when the correlation coefficient is close to the admissible boundaries. For the remainder, we use the square root of the inverse of the Fisher information for the scaling term of the score for it appears to be slightly preferred to other scaling choices as we found in the Monte Carlo study. Consequently, we evaluate correlation coefficient at $\rho = \tilde{\rho}_t$ through the reverse transformation of (4.19), where we assume that $\tilde{\rho}_t$ is available as it is computed in the previous time period. To model long-run dependence of time-varying correlation, we apply the long memory specification

$$(1 - L)^d(\tilde{g}_{t+1} - \mu) = \beta(1 - L)^d(\tilde{g}_t - \mu) + \alpha\tilde{\varepsilon}_t \qquad t = 1, \ldots, n, \tag{4.26}$$

where

$$\tilde{\varepsilon}_t = \tilde{\mathcal{I}}_t^{-1/2}\tilde{\nabla}_t, \qquad \tilde{\nabla}_t = \nabla_t|_{g=\tilde{g}_t}, \qquad \tilde{\mathcal{I}}_t = \mathcal{I}_t|_{g=\tilde{g}_t}, \tag{4.27}$$

with the score and Fisher information defined in (4.23) and (4.25) respectively.

4.2.5 Estimation of models

The common approach to estimate the parameters of the models within the observation-driven class of models is to directly maximize the associated log-likelihood function. The

copula approach is computationally convenient as it allows the marginal models be estimated separately from their dependence structure. Firstly, we estimate each marginal series independently considering a univariate volatility specification as discussed in Section 4.2.2. To make the notation clear, we add to subscript an integer i to denote the ith asset, $i = 1, 2$. The log-likelihood function of model for marginal based on the t density can be written as

$$\begin{aligned}\mathcal{L}_i^{T,m} &= n\left[\log\Gamma\left(\frac{\nu_i+1}{2}\right) - \log\Gamma\left(\frac{\nu_i}{2}\right) - \frac{1}{2}\log(\pi) - \frac{1}{2}\log(\nu_i-2)\right] \\ &\quad - \sum_{t=1}^{n}\log(\sigma_{it}) - \frac{\nu_i+1}{2}\sum_{t=1}^{n}\log\left(1+\frac{y_{it}^2}{(\nu_i-2)\sigma_{it}^2}\right).\end{aligned}$$

Estimation results of the marginal models provide us with approximate uniform margins obtained from standardized residuals $u_{it} = T_{\nu_i}(y_{it}/\sigma_{it})$. The obtained margins are used in the second step of the estimation procedure. Assuming the t copula density for the dependence structure, the log-likelihood is given by

$$\begin{aligned}\mathcal{L}^T &= n\left[\log\Gamma\left(\frac{\kappa+2}{2}\right) + \log\Gamma\left(\frac{\kappa}{2}\right) - 2\log\Gamma\left(\frac{\kappa+1}{2}\right)\right] \\ &- \frac{1}{2}\sum_{t=1}^{n}\log(1-\rho_t^2) - \frac{\kappa+2}{2}\sum_{t=1}^{n}\log\left(1+\frac{1}{\kappa(1-\rho_t^2)}\left(x_{1t}^2+x_{2t}^2-2\rho_t x_{1t}x_{2t}\right)\right) \\ &+ \frac{\kappa+1}{2}\sum_{t=1}^{n}\sum_{i=1}^{2}\log\left(1+x_{it}^2/\kappa\right),\end{aligned}$$

with $x_{it} = T_\kappa^{-1}(u_{it})$ and with the updating scheme for ρ_t as discussed in Section 4.2.4. The log-likelihood for the bivariate Gaussian copula density (4.18) can be written as

$$\mathcal{L}^G = -\frac{1}{2}\sum_{t=1}^{n}\log(1-\rho_t^2) - \frac{1}{2}\sum_{t=1}^{n}\left(\frac{1}{1-\rho_t^2}\left(x_{1t}^2+x_{2t}^2-2\rho_t x_{1t}x_{2t}\right) - \left(x_{1t}^2+x_{2t}^2\right)\right),$$

where now $x_{it} = \Phi^{-1}(u_{it})$.

The maximum likelihood estimation for the fractionally integrated class of models requires the truncation of the infinite distributed lags in (4.5). Since, we construct the fractionally integrated models to capture long memory features in the temporal dependence in the processes for volatility or dependence, we rather not truncate the polynomial (4.5) at too short lag. Given that financial time series are typically long enough, we follow Baillie et al. (1996) by estimating models using fixed truncation at lag 1000. We need also to set initial conditions to start up the recursions. With respect to parameter estimation for fractionally integrated GARCH models, Bollerslev and Mikkelsen (1996) argue that initialization has a negligible effect on the parameter estimates when the sample size is

large enough. They conveniently suggest to put innovation term to zero for the pre-sample values, while the values of h_t and of g_t for $t = \ldots, 0, 1$ can be fixed at unconditional sample moments. We report our simulation and empirical results using similar approach. The model parameters are collected in the parameter vector ψ and we maximize the log-likelihood function with respect to ψ. For example, we have $\psi = (\mu, \alpha, \beta, \kappa, d)'$ for the long memory model (4.26) with time-varying correlations and based on the bivariate t copula density. Maximization is carried out by a quasi-Newton numerical optimization procedure for which initial values of ψ are determined from a basic grid search.

4.3 Monte Carlo Study

4.3.1 Simulation design

In this part we carry out a Monte Carlo study to investigate the performance of some competing models for a bivariate conditional dependence under a variety of correlation dynamics. Design of our Monte Carlo study is similar to the study conducted by Engle (2002a) and Hafner and Manner (2010). We simulate a series of $n = 5000$ observations from the bivariate t copula with $\kappa \in \{2, 5, 15\}$ degrees of freedom by adopting various time-varying correlation ρ_t dynamics. For a decreasing number of the degrees of freedom κ, the t copula has more extreme realizations which may affect inference on correlation coefficient if that feature is not explicitly taken into account. Our goal is to contrast performance of some competing models and we also aim to study the ability of modeling the long-run dependence based on our modeling framework. The following eight processes are here considered:

1. *Constant* $\rho_t = 0.9$;
2. *Sine* $\rho_t = 0.5 + 0.4\cos(2\pi t/1000)$;
3. *Fast Sine* $\rho_t = 0.5 + 0.4\cos(2\pi t/100)$;
4. *Step* $\rho_t = 0.9 - 0.5(t > 2500)$;
5. *Ramp* $\rho_t = \mathrm{mod}(t/1000)$;
6. *ARMA* $(1 - 0.99L)g_t = 0.01 + 0.05\xi_t$;
7. *ARFIMA* $(1 - 0.90L)(1 - L)^{0.45}(g_t - 1) = 0.05\xi_t$;
8. *Random Walk* $(1 - L)g_t = 0.025\xi_t$;

where $\xi_t \sim \mathcal{N}(0, 1)$ is stochastic innovation, and where for stochastic processes 6 to 8 we apply the reverse of the transformation (4.19) to ensure ρ_t is well defined on its support at all times. Figures for the deterministic processes 1 to 5 can be found in Engle (2002a). The considered processes reflect a diversity of dynamics with sudden breaks, smooth changes, regime type behavior, periods of constancy, different speed of mean-reversion and different structure of temporal dependence in the underlying dynamics.

We simulate $M = 1000$ bivariate series for each of the eight processes for time-varying correlation. For convenience, we refer to our model presented in Section 4.2.4 to as the Fractionally Integrated Generalized Autoregressive Score model, or FIGAS, which corresponds to the original paper of Creal et al. (2008). Since the long memory formulation (4.21) nests conventional short as well as infinite memory models as special cases, we compare the following (FI)GAS specifications: short memory GAS$(1, d \equiv 0, 1)$, infinite memory IGAS$(\beta \equiv 0, d \equiv 1, 1)$, and long memory FIGAS$(1, d, 1)$. To contrast our models, we benchmark them against the model recently proposed by Dias and Embrechts (2010), where the update for the transformed correlation g_t has the following form

$$g_{t+1} = \delta + \beta g_t + \alpha \operatorname{sgn}(x_{1t}\, x_{2t}) |x_{1t}\, x_{2t}|^{1/2}, \tag{4.28}$$

where $\operatorname{sgn}(\cdot)$ is the signum function. Dias and Embrechts (2010) find that their specification outperforms some other competing models proposed in the literature. The use of the square root of the product $|x_{1t}\, x_{2t}|$ dampens some more extreme observations; one could also use $|x_{1t}\, x_{2t}|^{\lambda}$ where λ would be jointly estimated but we do not pursue this idea here. Dias and Embrechts (2010) refer to the updating scheme (4.28) as to the Fisher$(1, 1)$ dynamics.

For our purpose, we consider also some baseline long memory model,

$$(1 - L)^d (g_{t+1} - \mu) = \beta (1 - L)^d (g_t - \mu) + \alpha \left(\frac{1}{H} \sum_{j=1}^{H} x_{1t-j+1} x_{2t-j+1} - \rho_t \right), \tag{4.29}$$

where $H = 10$ is the smoothing parameter in a similar way as proposed by Patton (2002, 2006). This model specification weights equally all observations and does not exploit features of the t density. We refer to (4.29) as to the FIBase$(1, 1)$ model.

Given the simulated data, we estimate the models by maximizing the log-likelihood function as discussed in Section 4.2.5. We next relate estimated correlation $\widehat{\rho}_t$ to the true known correlation ρ_t, and we report two measures of accuracy: the mean absolute error (MAE) and the mean squared error (MSE),

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^{n} |\rho_t - \widehat{\rho}_t| \qquad \text{and} \qquad \text{MSE} = \frac{1}{n} \sum_{t=1}^{n} (\rho_t - \widehat{\rho}_t)^2, \tag{4.30}$$

that are averaged over the number of repetitions $M = 1000$.

4.3.2 Simulation results

Table 4.1 presents simulation results. The obtained MAE and MSE statistics are reported relative to the smallest value within the model specification taken for the true dynamics of ρ_t in the data generating process, and the smallest values are presented in bold font. For

Table 4.1: *Simulation results*

	Constant	*Sine*	*Fast Sine*	*Step*	*Ramp*	*ARMA*	*ARFIMA*	*Rand Walk*
				$\kappa = 2$				
MAE								
GAS	1.1193	1.0442	1.0084	**1.0000**	1.0005	**1.0000**	1.0043	**1.0000**
IGAS	**1.0000**	1.0579	1.1171	1.0983	**1.0000**	1.0638	1.0487	1.0316
FIGAS	1.0803	**1.0000**	**1.0000**	1.1431	1.0143	1.0059	**1.0000**	1.0154
Fisher	1.3548	1.7411	1.3264	2.2583	1.3220	1.1095	1.1116	1.2402
FIBase	1.9438	2.8418	1.9488	3.3515	2.0036	1.3600	1.5003	2.4370
MSE								
GAS	**1.0000**	1.0785	1.0100	**1.0000**	**1.0000**	**1.0000**	1.0090	**1.0000**
IGAS	1.1004	1.1147	1.3172	1.0903	1.0760	1.1191	1.0957	1.1103
FIGAS	1.1085	**1.0000**	**1.0000**	1.1184	1.0299	1.0109	**1.0000**	1.0793
Fisher	1.8491	3.2993	1.6754	4.5628	1.7101	1.2492	1.2175	1.5865
FIBase	4.4886	6.5415	3.5528	7.4602	3.5494	3.6131	2.4562	5.0365
				$\kappa = 5$				
MAE								
GAS	1.1301	1.0333	1.0081	**1.0000**	**1.0000**	**1.0000**	1.0035	**1.0000**
IGAS	1.3113	1.0726	1.1373	1.1129	1.0065	1.0794	1.0529	1.0383
FIGAS	**1.0000**	**1.0000**	**1.0000**	1.1252	1.0082	1.0107	**1.0000**	1.0316
Fisher	1.2891	1.4423	1.2208	1.7350	1.1572	1.0835	1.0920	1.2095
FIBase	1.7225	1.5756	1.4291	1.9037	1.2396	1.1415	1.1444	1.2639
MSE								
GAS	1.1737	1.0498	1.0058	**1.0000**	**1.0000**	**1.0000**	1.0062	**1.0000**
IGAS	1.2648	1.1439	1.3668	1.1110	1.0779	1.1659	1.1123	1.1365
FIGAS	**1.0000**	**1.0000**	**1.0000**	1.1024	1.0263	1.0217	**1.0000**	1.1269
Fisher	1.4266	1.8976	1.3550	2.3669	1.2916	1.1644	1.1713	1.4932
FIBase	2.0972	2.1497	1.7083	2.5462	1.4179	1.2876	1.2792	1.5864
				$\kappa = 15$				
MAE								
GAS	1.3411	1.0313	1.0090	**1.0000**	**1.0000**	**1.0000**	1.0047	**1.0000**
IGAS	1.1192	1.0755	1.1339	1.1217	1.0004	1.0844	1.0585	1.0445
FIGAS	**1.0000**	**1.0000**	**1.0000**	1.1275	1.0116	1.0054	**1.0000**	1.0330
Fisher	1.0469	1.4330	1.2312	1.7496	1.1478	1.0858	1.0941	1.2153
FIBase	1.3092	1.5143	1.2148	1.8601	1.2432	1.1232	1.1350	1.2790
MSE								
GAS	1.3601	1.0433	1.0046	**1.0000**	**1.0000**	**1.0000**	1.0089	**1.0000**
IGAS	1.1501	1.1468	1.3714	1.1275	1.0623	1.1816	1.1277	1.1523
FIGAS	**1.0000**	**1.0000**	**1.0000**	1.1147	1.0357	1.0115	**1.0000**	1.1335
Fisher	1.0890	1.8743	1.3409	2.4567	1.3436	1.1697	1.1778	1.5292
FIBase	1.3769	1.9972	1.3782	2.4635	1.4425	1.2489	1.2588	1.6194

example, for $\kappa = 2$ and *step* specification, the IGAS model appears to outperform other models.

The results can be summarized as follows. We observe that (FI)GAS models perform reasonably well and stable regardless of the number of degrees of freedom κ. In contrast, models (4.28) and (4.29) perform considerably worse when low number of degrees of freedom $\kappa = 2$ is applied. The t copula with a low number of degrees of freedom is often said to have a star shape implying occurrence of more extreme realizations, and these may affect the update of correlation if not being downweighted appropriately. Only when κ is higher, performance of models (4.28) and (4.29) corresponds closer to the (FI)GAS models. In general, the short memory GAS specification appears to preferred in the setups when the

correlation process is subject to either sudden breaks or abrupt changes (*step* and *ramp*), and when the true correlation process is a short memory $ARMA$ process. On the other hand, the long memory FIGAS specification appears to be preferred when the correlation process changes smoothly over time (*sine* and $fast\ sine$), and when the true correlation process is a long memory ARFIMA process. This finding illustrates that the introduction of fractional integration is effective and it shows that the FIGAS model identifies correctly long-run dependence in the correlation dynamics. The performance of the (FI)GAS models does not vary with κ which is a satisfactory result. When the correlation follows a random walk specification, the GAS performs best with most β estimates being statistically not different from a unity. It follows from the Table 4.1 that these results in terms of MAE and MSE measures lead to similar conclusions. The FIGAS formulation is preferred to other specifications when the temporal dependence in the correlation process has a slow rate of decay or when the true process changes smoothly over time.

By considering the dynamic t copula when the true generating process is the dynamic Gaussian copula, the estimation procedure of the (FI)GAS specification for the transformed correlation often does not converge as the log-likelihood function is increasing with the number of degrees of freedom κ. Alternatively, when estimating the dynamic Gaussian copula when the true generating process is based on the dynamic t copula, the estimation procedure is often unstable as the Gaussian based update for correlation implies $\pi_t = 1$ in (4.23). It therefore lacks robustness against outlying observations and it can lead to an almost singular correlation matrix. Similar findings apply to the dynamics of (4.28) and (4.29). These insights may prove helpful in dealing with parameter estimation for empirical models.

4.4 Empirical Illustrations

4.4.1 Data and proxies for latent processes

The data we analyze includes some four Dow Jones Industrial Average components: American Express (NYSE ticker symbol: AXP), General Electric (GE), The Coca-Cola Company (KO) and Procter & Gamble (PG). We use the tick data on consolidated trades extracted from the Trade and Quote (TAQ) database through the Wharton Research Data Services (WRDS) system. The data is taken from the NYSE core trading session between the hours of 9:30am to 4:00pm Eastern Time. The sample period runs from January 4, 1993 to May 28, 2010 with a total of $n = 4385$ trading days for all four equities. We carried out the cleaning procedures on the raw trade data applying the methodology discussed by Barndorff-Nielsen et al. (2009). We put all data to the calendar time sampling using the previous tick method and by aggregating records with the same second precision time-stamp into one observation using the median price; see Hansen and Lunde (2006). Our

study considers modeling of daily returns that we define as the difference between the logarithm of the close and the logarithm of the open price. Price increments sampled at the 5 minute time intervals are used to compute realized measures which serve to benchmark our models. At 5 minute sampling frequency the impact of microstructure noise is benign and can be ignored in our practical application. To be consistent with the definition of daily returns, we skip the overnight returns in the computations of realized measures.

The availability of intraday high-frequency data facilitates computation of the realized measures like realized volatility or covariance. Following Andersen and Bollerslev (1998a) we use realized measures as proxies for true latent processes, so that we can further evaluate the model-based volatilities and correlations across different models. For our purpose, we use the threshold realized volatility proposed by Mancini (2009) defined as

$$TRV_N^k = \sum_{i=1}^{N} (\Delta X_{t_i}^k)^2 \mathbf{1}_{\{(\Delta X_{t_i}^k)^2 \leq r_N\}}, \quad \Delta X_{t_i} = X_{t_i} - X_{t_{i-1}}, \tag{4.31}$$

where $X_{t_i}^k$ denotes the log price of asset k at the observation time t_i, $i = 0, \ldots, N$; here $N = 78$ denotes the number of intraday returns sampled at the $5min$ frequency; function $\mathbf{1}_{\{(\Delta X_{t_i}^k)^2 \leq r_N\}}$ takes value one if $(\Delta X_{t_i}^k)^2 \leq r_N$ and zero else, where r_N denotes the threshold function such that $r_N = N^{-0.99}$ following the suggestion of Mancini (2009). The idea behind the threshold estimator (4.31) is built on the statistical properties of the Brownian motion whose variation is proportional to the time step $N^{-1/2}$. A price increment whose variation is larger than that of the Brownian variation might contain a jump component, thus the use of the threshold function r_N makes effectively jumps be discarded from the estimation. The threshold realized correlation between assets k and l proposed by Mancini and Gobbi (2010) is defined as

$$TRCorr_N^{k,l} = \frac{\sum_{i=1}^{N} \Delta X_{t_i}^k \mathbf{1}_{\{(\Delta X_{t_i}^k)^2 \leq r_N\}} \Delta X_{t_i}^l \mathbf{1}_{\{(\Delta X_{t_i}^l)^2 \leq r_N\}}}{\sqrt{\sum_{i=1}^{N} (\Delta X_{t_i}^k)^2 \mathbf{1}_{\{(\Delta X_{t_i}^k)^2 \leq r_N\}} \sum_{i=1}^{N} (\Delta X_{t_i}^l)^2 \mathbf{1}_{\{(\Delta X_{t_i}^l)^2 \leq r_N\}}}}. \tag{4.32}$$

The threshold estimators measure variation attributable to the continuous part of the price process. Their use as proxies for latent processes might provide a further insight into performance of models under study in the context of the sensitivity to some possible jumps and/or outliers in the equity data. Without truncation classical realized volatility (RV) and realized correlation ($RCorr$) measures are retrieved as introduced by Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002). We compute the threshold realized variance and threshold realized correlation for each day in our sample and the equities considered.

4.4.2 Conditional volatility of returns on equity

In the first part of our empirical illustration we perform a comparative study by fitting some competing models to the marginals as discussed in Section 4.2.2. In addition, within each particular model we consider both short as well as long memory specification. As a benchmark framework, we select the well-known (FIE)GARCH class of models and we consider the following eight formulations: GARCH, FIGARCH, EGARCH, FIEGARCH, GAS, FIGAS, GAS+L(everage) and FIGAS+L(everage). All models are assumed to have the most parsimonious specification; see the note at the bottom of Table 4.B.I for their detailed description.

Table 4.B.I presents parameter estimates and associated standard errors. As a starting point, we provide the estimates of the fractional order of integration d of the (threshold) realized volatility using the log-periodogram regression method of Geweke and Porter-Hudak (1983), or the so-called GPH estimate. It is seen from the two most left columns that estimates are in between 0 and 1 suggesting that all financial volatilities have long range dependence and are fractionally integrated processes; see also Maasoumi and McAleer (2008) for more discussion. The persistence of the volatility processes can be also inferred from the estimates of the autoregressive coefficient β, or in the case of the GARCH model from $\alpha+\beta$, that are in the range $0.9918(0.0027) \leq \widehat{\beta} \leq 0.9969(0.0018)$ for the short memory models under study. Most of the empirical studies based on the ARCH family of models have found the volatility process to be highly persistent and possibly not covariance-stationary, see Xekalaki and Degiannakis (2010). It is also often found that volatility appears to increase proportionally more for negative than for positive shocks. The (FI)EGARCH and (FI)GAS+L models are designed to accommodate for such asymmetric relation, and it is seen that estimates of γ are statistically different than zero at the conventional levels suggesting the importance of the leverage effect for all four equities. The estimates of ν are in the range $7.4017(0.7273) \leq \widehat{\nu} \leq 12.3729(1.8834)$ and the confidence intervals cover some common part of the parameter space.

The short memory models seem too restrictive in view of the GPH estimates, thus we report also parameter estimates for the long memory alternatives. It can be seen that for all four equities and for all four models estimates of d are statistically different from both zero and one at the conventional significance levels and are in the range $0.4392(0.0294) \leq \widehat{d} \leq 0.7427(0.0395)$. This finding implies that innovations have a long-run effect on the future volatility. The AIC and SIC criteria suggest that in most cases the volatility is best described by the non-stationary but mean-reverting fractionally integrated process. The obtained results conform with the findings of Baillie et al. (1996) and Bollerslev and Mikkelsen (1996) who applied FI(E)GARCH modeling tools. Within the models, the degree of long memory in volatilities is similar for all four equities which corresponds to the findings of Pafka and Mátyás (2001) and Conrad et al. (2010).

Figure 4.3: *Conditional volatility of returns on the PG equity*

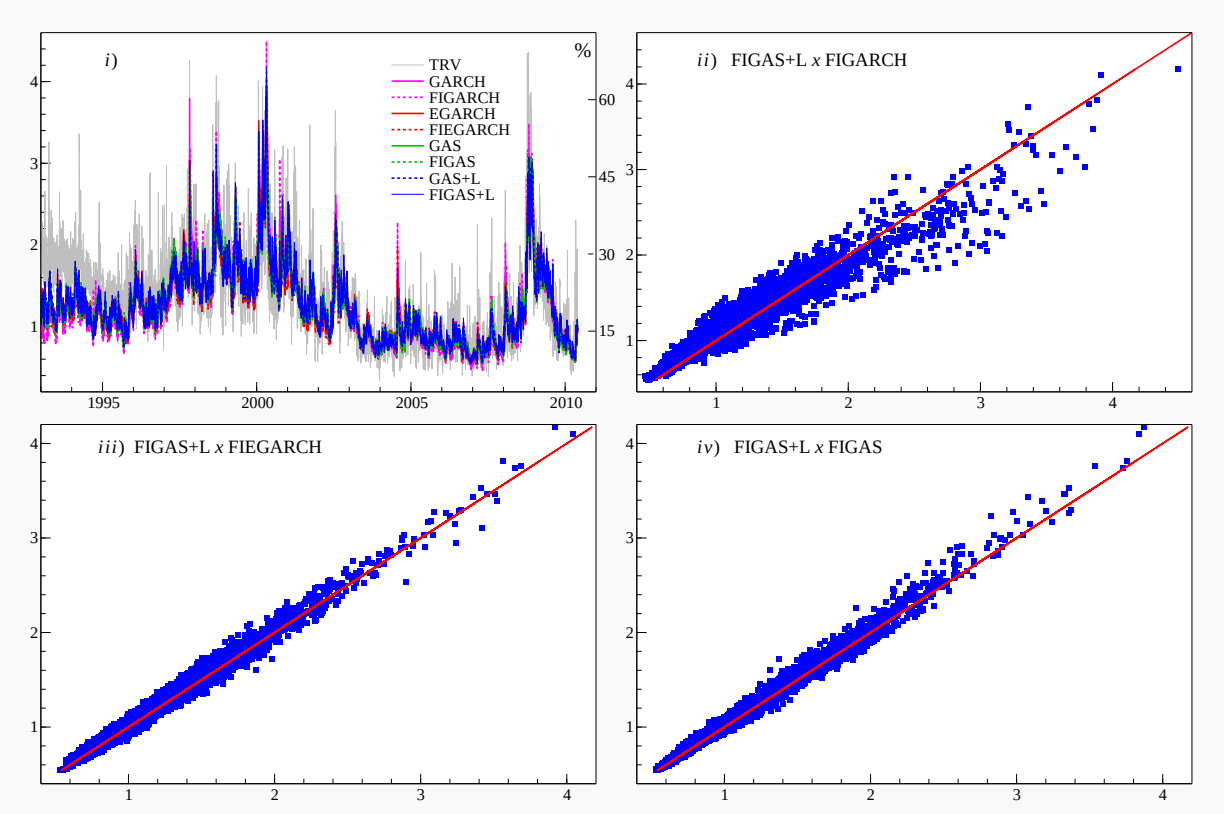
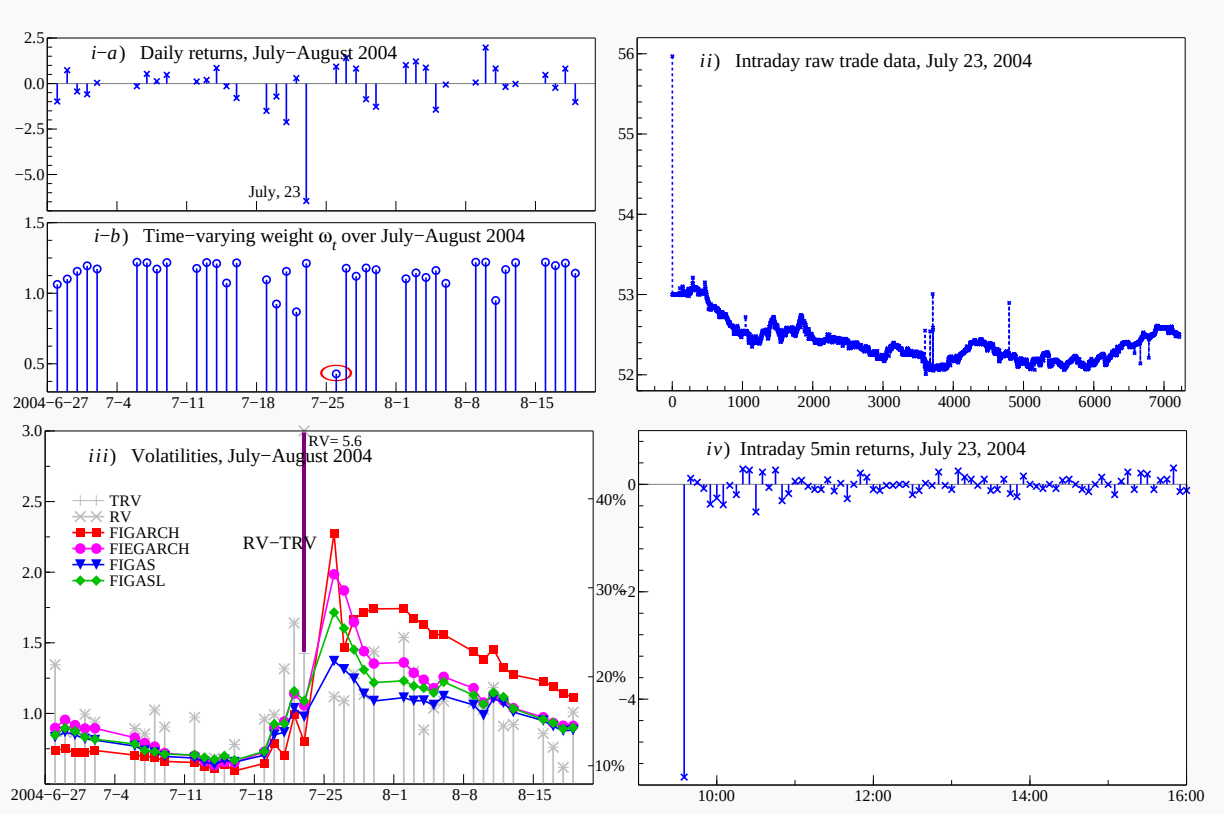


Figure 4.3 gives model-based volatilities, to be specific we plot σ_t and $\sqrt{TRV_t}$. Panel i) presents model-based volatilities plotted along with the threshold realized volatility over the sample period for the PG equity. It is seen that the realized measure is much more jagged than the model-based volatilities, though in general similar dynamics can be recognized. The time series of estimated volatilities indicate volatile and tranquil periods with the average annualized volatility (right scale) of around 25%. Panels ii) to iv) present FIGAS+L volatility plotted against other three long memory models under study. A substantial difference is visible when contrasting the FIGAS+L and FIGARCH volatillities, while the model-based volatilities from the FIGAS+L, FIGAS and FIEGARCH models are rather similar. However, there are noteworthy differences around considerable movements in the equity prices.

One such case study is illustrated in Figure 4.4. On July 23, 2004 the daily return on the PG equity was around -6.5% as seen in panel i-a) with the open price 55.97\$ and the close price 52.49\$ as seen in panel ii). This daily price increment was much larger (in absolute terms) than in the adjacent days, see panel i). Having high-frequency intraday data we observe that after the first trade, the rest of transactions over a course

Figure 4.4: *Case study for the PG equity*



of this day was recorded at much lower price level of around 52.50-53.00, see panel *ii*) of Figure 4.4. Consequently, the first computed intraday return and daily return are both substantially large and we see in panel *iii*) it has a noticeable influence on the estimated volatilities. Firstly, we perceive a substantial divergence between RV and TRV on this particular day. Secondly, the model-based volatilties rise instantly and we observe that it takes some time to return to the previous level (and to TRV) after being hit by this rather extreme innovation, which is more evident in the FI(E)GARCH models than in the FIGAS alternative. The FIGAS updating scheme downweights this extreme observation via the weight ω_t as defined in (4.9), and the resulting volatility estimate remains more robust; see panel *i-b*) of Figure 4.4. The volatility dynamics in the FIGAS+L specification is designed to respond proportionally more to negative shocks than the baseline FIGAS specification, therefore we note a higher impact. However, the FIGAS+L volatility remains more robust than the FI(E)GARCH volatilities.

We conclude the first part of our empirical study by commenting on the last two columns of Table 4.B.I that relate the model-based volatilities to high-frequency realized measures using the MAE statistics. We find that the FIGAS+L model for financial volatility for

these equities appears to be preferred when judged by this criteria.

4.4.3 Conditional dependence between returns on equities

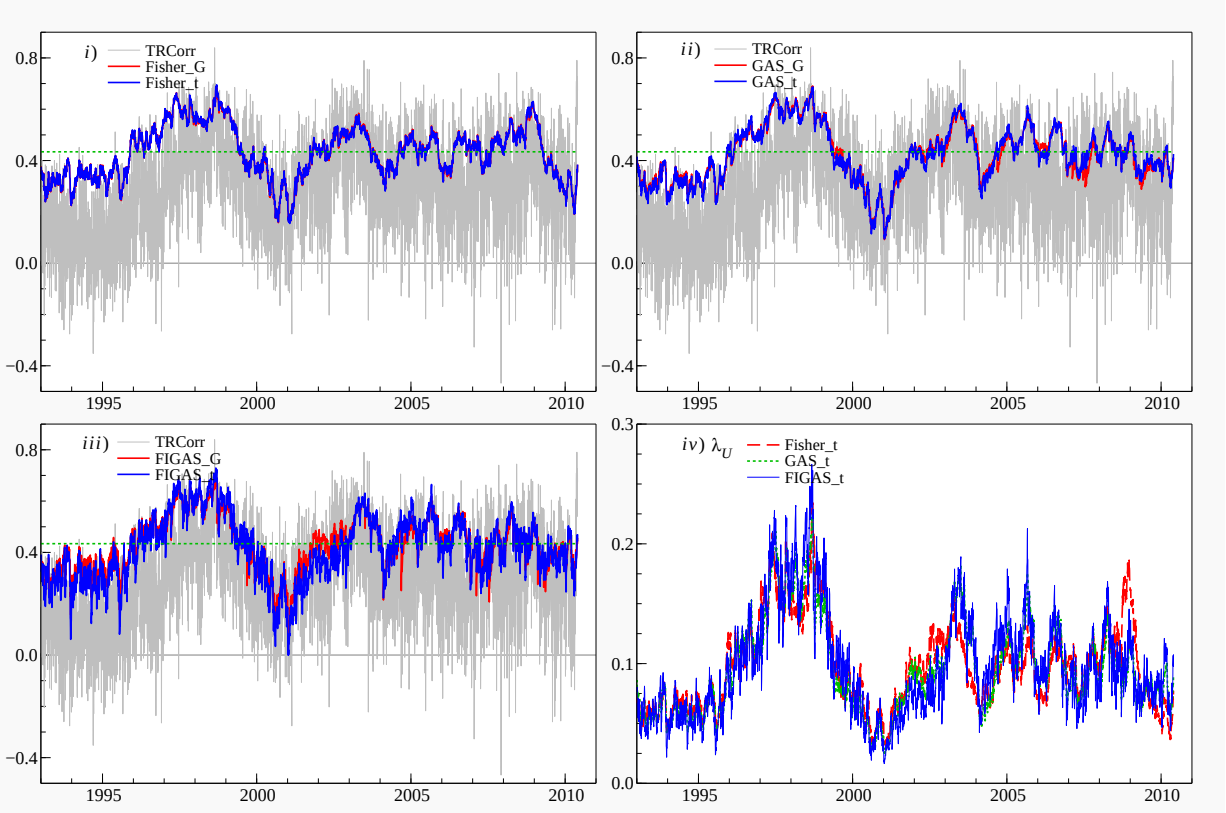
In the second part of our empirical study we consider modeling of dependence structure as discussed in Section 4.2.4. Using the estimates of individual volatilities, we transform the standardized returns into uniform variables using the probability integral transform. We estimate a set of competing models for each possible pair within the four equities under consideration.

Table 4.B.II presents maximum likelihood parameter estimates and associated standard errors. We estimate the short as well as long memory models for dependence for the bivariate Gaussian and t copula densities. For a preliminary insight, we provide the GPH estimates of the fractional order of integration d of the (threshold) realized correlation. The obtained GPH estimates are near 0.5 value and they hint at long memory features in the correlation processes, likewise in covariance process. We observe also that the GPH estimates for realized correlations are slightly smaller than the GPH estimates of realized volatilities, but still the realized correlations appear to be fractionally integrated processes. Within the model-based framework, the persistence of the correlation processes can be inferred from the estimates of the autoregressive coefficient β that are in the range $0.9037(0.0322) \leq \widehat{\beta} \leq 0.9969(0.0025)$ for the short memory models. The value of the lower bound of the range of β estimates for correlations is smaller than in the case of volatilities, which might suggest that the degree of memory in these processes may differ. The correlation process is found generally to be persistent and this finding is in accordance with the previous studies based on other short memory specifications. The empirical results of Dias and Embrechts (2010) indicate that high persistence of dependence process in dynamic copulas is a common feature at different (reasonable) sampling frequencies of financial returns. Hafner and Manner (2010) report that in some copulas a dependence process is even more likely to be an integrated process rather than stationary as the persistence parameter in their short memory specification attains the upper bound of the admissible parameter space for stationary processes.

To obtain greater flexibility we allow for a fractional order of integration and report estimation results for the FIGAS based long memory specification. The estimates of the fractional order of integration d are in the range $-0.0053(0.0184) \leq \widehat{d} \leq 0.7138(0.0768)$ in case of the FIGAS model. In four pairs we find evidence of long range dependence in the correlation dynamics as judged by the maximum likelihood estimates of fractional order of integration being different from zero at the conventional levels; similar findings arise from the estimates of d in the FIBase model as defined in (4.29). In case of the pair AXP/GE the estimate of d being statistically different from 0.5 suggests that the dependence might not be covariance-stationary but still mean-reverting process. In two pairs AXP/GE and

GE/KO the degrees of memory (as judged by estimates of d) in volatilities and correlation processes appear to be comparable, while in other four pairs the degree of memory in the correlation process is lower than in the volatility processes. Our result for equities question the approach of Pafka and Mátyás (2001) who impose a common degree of long memory parameter for volatilities and covariances in their trivariate FIGARCH model for some three exchange rates series. The last two columns of Table 4.B.II relate the model-based and realized correlations using the MAE statistics. We observe that the dynamic t copula performs better than the dynamic Gaussian copula. Finally, we find that the long memory FIGAS model based on the bivariate t copula appears to be preferred when judged by the MAE criteria.

Figure 4.5: *Conditional dependence between returns on GE and KO equities*



Note: Dotted line is the estimate of ρ in the static bivariate t copula.

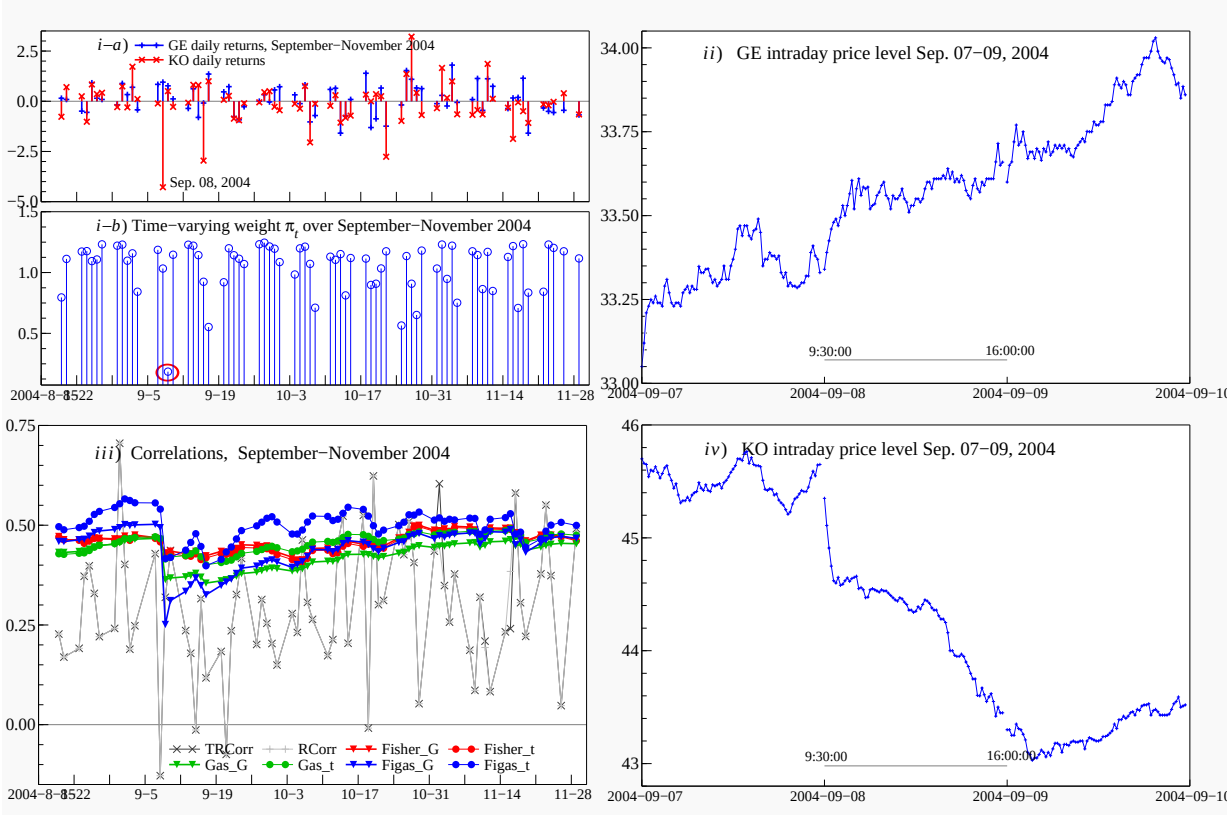
Figure 4.5 presents the time-varying correlations plotted along with the threshold realized correlation for one pair of equities. The realized correlation is much more jagged than the model-based correlations. The GE/KO dependence oscilattes around the value obtained from the static copula (long-run mean) without any significant breaks. The t copula density generalizes the Gaussian copula density by allowing for non-zero dependence in the tails of the distribution. Both parameters of the t copula, i.e. correlation ρ

and the degrees of freedom η, determine the tail dependence. Since the t copula density is symmetric the upper and lower tail dependence are equal,

$$\lambda_L = \lambda_U = 2t_{\eta+1}\left(- \sqrt{\eta+1}\frac{\sqrt{1-\rho}}{\sqrt{1+\rho}}\right) \in [0,1]. \tag{4.33}$$

Panel iv) of Figure 4.5 presents the implied time-varying tail dependence. It appears that the highest dependence between extreme events for this pair was around 1997-1998.

Figure 4.6: *Case study for the GE/KO equities*



The resulting model-based correlations are somewhat similar as seen in panels i)-iii) of Figure 4.5. However, the models behave noticeably different when some more extreme observations happen to occur in the sample. One such example is illustrated in Figure 4.6 for the GE/KO pair. Panel i-a) of Figure 4.6 indicates that on September 8, 2004 the daily return on the GE equity was quite positive, while return on the KO equity was substantially negative, around -4.5%. To provide some further insight, panels ii) and iv) present intraday price level (sampled at the $5min$ frequency) for three days and it is seen that on September 8 the two equity prices clearly go in the opposite directions. This occurence has its effect on

the resulting model-based correlations as seen in panel *iii*). The correlations soar sharply downwards and afterwards they return slowly to the previous level. It is seen in panel *i-b*) that (FI)GAS models based on t copula ((FI)GAS-t) adapt the correlation update via π_t as defined in (4.24) for more extreme observations and consequently the estimates are more robust than corresponding models based on the Gaussian copula. In particular, the future values of the (FI)GAS based Gaussian correlations appear to be too heavily affected by the occurence of somewhat more extreme return realizations.

4.5 Summary and Conclusions

We have proposed new fractionally integrated formulation for the purpose of modeling long memory dynamics in volatility and dependence. The introduced models add some flexibility in modeling short and long range dependence as they allow innovations to dissipate at slow hyperbolic rate. The allowance for a long lasting impact of innovations for future values calls however forth a treatment of possible outliers or extreme observations in the empirical data. Since conditional heavy tails and dependence between extreme events are both relevant features of financial data, our modeling of the joint distribution is based on the copula approach. The model for marginals is based on the heavy-tailed t density, while the model for dependence is based on the bivariate t copula density. Dynamics for volatility and dependence are constructed using the scores of the t density. In consequence, innovations are pre-multiplied with time-dependent weights which downweight more extreme observations making the updating schemes be more robust. We have illustrated in the empirical applications that the robustness property is the key distinguishing feature of our modeling framework relative to some other competing alternatives, both when modeling volatility and dependence. We have found that the degree of long memory in volatilities is rather similar across equities, while the degree of long memory in correlations across equity pairs may be substantially different and different than the degree of long memory in the volatilities. We have evaluated our models using high-frequency realized measures, and we have found them performing well when compared to some other competing models.

4.A Appendix

4.A.1 Autocorrelation function

Let h_t process denote time-varying parameter of interest modeled via the general long memory formulation (4.4) as given by

$$B(L)(1-L)^d(h_{t+1}-\mu) = A(L)\eta_t, \tag{4A.1}$$

where η_t is martingale difference assumed to be mean zero and with time-invariant finite variance σ_η^2 and where polynomials $B(L)$

$$\begin{aligned} B(L) &= 1-\beta_1 L-\beta_2 L^2-\ldots-\beta_p L^p, & \text{(4A.2)}\\ A(L) &= \alpha_0+\alpha_1 L+\alpha_2 L^2+\ldots+\alpha_{q-1}L^{q-1}. & \text{(4A.3)} \end{aligned}$$

The algorithm for the computation of the variance-autocovariance function of the general ARFIMA process is derived by Sowell (1992), and it can be adapted for the process (4A.1). Let the autoregressive polynomial $B(L)$ be written by using its roots as follows

$$B(L) = \prod_{j=1}^{p}(1-\widetilde{\beta}_j L). \tag{4A.4}$$

The autocovariance function of (4A.1) is given by

$$\gamma_k = \sigma_\eta^2 \sum_{i=-q+1}^{q-1}\sum_{j=1}^{p}\varphi_i\zeta_j G(d, p+i-k, \widetilde{\beta}_j), \tag{4A.5}$$

where

$$\varphi_i = \sum_{h=|i|}^{q-1}\alpha_h\alpha_{h-i}, \qquad \zeta_j = \left[\widetilde{\beta}_j\prod_{i=1}^{p}(1-\widetilde{\beta}_i\widetilde{\beta}_j)\prod_{m\neq j}^{p}(\widetilde{\beta}_j-\widetilde{\beta}_m)\right]^{-1}, \tag{4A.6}$$

with $\varphi_0=\alpha_0^2$; function G is defined as

$$\begin{aligned} G(d,h,\widetilde{\beta}) = \frac{\Gamma(1-2d)}{\Gamma(1-d)\Gamma(d)}\frac{\Gamma(d+h)}{\Gamma(1-d+h)}\Big[&\widetilde{\beta}^{2p}H(d+h,1,1-d+h,\widetilde{\beta})\\ &+H(d-h,1,1-d-h,\widetilde{\beta})-1\Big], \end{aligned} \tag{4A.7}$$

where $H(a,b,c,x)$ denotes the hypergeometric function

$$H(a,b,c,x) = \sum_{i=0}^{\infty}\frac{(a)_i(b)_i}{(c)_i}\frac{x^i}{i!}, \tag{4A.8}$$

and $(a)_i$, likewise $(b)_i$ and $(c)_i$, denotes the Pochhammer's symbol $(a)_i = \prod_{j=0}^{i-1}(a+j)$ with $(a)_0 = 1$, such that (4A.8) can be written as

$$H(a,b,c,x) = 1 + \frac{a\,b}{c}x + \frac{a(a+1)\,b(b+1)}{c(c+1)}\frac{x^2}{2} + \ldots. \tag{4A.9}$$

Using (4A.5) the autocorrelation at lag k, i.e. γ_k/γ_0, is thus given by

$$\begin{aligned} c_k = \frac{\Gamma(d+p+i-k)}{\Gamma(d+p+i)}\frac{\Gamma(1-d+p+i)}{\Gamma(1-d+p+i+k)} \\ \times \left\{ \frac{\sum_{i=-q+1}^{q-1}\sum_{j=1}^{p}\varphi_i\zeta_j G(d,p+i-k,\widetilde{\beta}_j)}{\sum_{i=-q+1}^{q-1}\sum_{j=1}^{p}\varphi_i\zeta_j G(d,p+i,\widetilde{\beta}_j)} \right\}, \end{aligned} \tag{4A.10}$$

and does not depend on σ_η^2.

4.A.2 *t* density

The density of the central multivariate t density (see Johnson and Kotz (1972)) is given by,

$$p(y|\mu,\Sigma,\vartheta) = \frac{\Gamma(\frac{\vartheta+k}{2})}{\Gamma(\frac{\vartheta}{2})(\pi\vartheta)^{k/2}|\Sigma|^{1/2}}\left[1+\vartheta^{-1}(y-\mu)'\Sigma^{-1}(y-\mu)\right]^{-\frac{\vartheta+k}{2}}, \tag{4A.11}$$

where $y = [y_1,\ldots,y_k]'$ is k dimensional random variable, μ is $k\times 1$ location vector, Σ is $k\times k$ positive semi-definite scale matrix, $\vartheta > 0$ is the number of degrees of freedom, and $\Gamma()$ is the Gamma function. For simplicity, we set $\mu = 0$. We are interested in score and expected information of the log density of (4A.11) with respect to φ where $\Sigma = \Sigma(\varphi)$.

Proposition 3. *Let $l(\varphi,\vartheta) = \log p(y|\mu,\Sigma,\vartheta)$. The score and expected information are given by:*

$$\frac{\partial l(\varphi,\vartheta)}{\partial\varphi} = -\frac{1}{2}tr\left(\Sigma^{-1}\frac{\partial\Sigma}{\partial\varphi}\right) + \frac{1}{2}\Omega\, y'\,\Sigma^{-1}\,\frac{\partial\Sigma}{\partial\varphi}\,\Sigma^{-1}\,y, \tag{4A.12}$$

$$\begin{aligned} \mathbb{E}\left[\left(\frac{\partial l(\varphi,\vartheta)}{\partial\varphi}\right)^2\right] &= \frac{\vartheta+k}{2(\vartheta+k+2)}\,tr\left(\Sigma^{-1}\frac{\partial\Sigma}{\partial\varphi}\Sigma^{-1}\frac{\partial\Sigma}{\partial\varphi}\right) \\ &\quad -\frac{1}{2(\vartheta+k+2)}\,tr\left(\Sigma^{-1}\frac{\partial\Sigma}{\partial\varphi}\right)tr\left(\Sigma^{-1}\frac{\partial\Sigma}{\partial\varphi}\right), \end{aligned} \tag{4A.13}$$

where $\Omega = \frac{\vartheta+k}{\vartheta+y'\Sigma^{-1}y}$.

Proof: See Lange et al. (1989). □

Case 1: We consider the case of Section 4.2.2. Let $\varphi = h \in \mathbb{R}$ and let $\vartheta^{-1}\Sigma = \sigma^2/(\nu-2)$, and $\sigma^2 = \exp h > 0$, $\nu > 2$, such that we consider variance rather than scale parameter. The results of (4.8) and (4.9) follow then immediately from Proposition 3.

Case 2: We consider the case of Section 4.2.4. Before we state the main result, we briefly summarize copula functions. A k-dimensional copula $C(u_1, \ldots, u_k)$ is a multivariate distribution function in the unit hypercube $[0,1]^k$ with uniform $u_i \sim U(0,1)$ marginal distributions, and $i = 1, \ldots, k$. It can be shown, Sklar's theorem, that every joint distribution, $F(x_1, \ldots, x_k)$, whose univariate marginals are given by $F_1(x_1), \ldots, F_k(x_k)$ can be written as

$$F(x_1, \ldots, x_k) = C\big(F_1(x_1), \ldots, F_k(x_k)\big), \tag{4A.14}$$

for a function C that is called a copula of F. If the marginal distributions are continuous, then there is a unique copula associated to the joint distribution F that is obtained as follows

$$C(u_1, \ldots, u_k) = F\big(F_1^{-1}(u_1), \ldots, F_k^{-1}(u_k)\big), \tag{4A.15}$$

where F_i^{-1} is the inverse cumulative distribution function of the ith marginal. Given a k-dimensional copula and k univariate marginal distributions, the density function of (4A.14) is given by

$$f(x_1, \ldots, x_k) = c\big(F_1(x_1), \ldots, F_k(x_k)\big) \prod_{i=1}^{k} f_i(x_i), \tag{4A.16}$$

and consequently the copula density reads

$$c\big(F_1(x_1), \ldots, F_k(x_k)\big) = \frac{f(x_1, \ldots, x_k)}{\prod_{i=1}^{k} f_i(x_i)}. \tag{4A.17}$$

It follows from (4A.17) that dependence structure in the copula density appears in the multivariate f density. The implication is that the score and expected information of the correlation matrix in the t copula density correspond to those from the multivariate t density.

In particular, let $\varphi = g$, $\vartheta = \kappa$, $k = 2$, and let

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \tag{4A.18}$$

with $\rho = \frac{\exp(g)-1}{\exp(g)+1} \in (-1, 1)$, and $g \in \mathbb{R}$. The results of (4.23) and (4.25) follow by applying these definitions to the expressions in Proposition 3.

4.B Maximum likelihood estimation results

Table 4.B.I: *Maximum likelihood estimates for conditional volatility of daily returns on equities*

Equity	d^{RV}	d^{TRV}	Model	θ	δ	α	γ	β	ν	d	AIC	SIC	MAE^{RV}	MAE^{TRV}
AXP	0.5492	0.5428	GARCH	0.0421	0.0143	0.0619	-	0.9341	9.6123	-	16970.70	17002.63	1.0000	1.0000
	(0.0284)	(0.0284)		(0.0205)	(0.0047)	(0.0082)		(0.0088)	(1.2037)					
			FIGARCH	0.0464	-	0.2499	-	0.7479	10.3376	0.5947	16996.78	17028.71	1.0188	1.0215
				(0.0203)		(0.0336)		(0.0432)	(1.2239)	(0.0561)				
			EGARCH	0.0215	0.0073	0.1343	-0.0448	0.9931	9.5393	-	16944.38	16982.70	0.9571	0.9528
				(0.0201)	(0.0027)	(0.0166)	(0.0087)	(0.0020)	(1.2744)					
			FIEGARCH	0.0495	-	0.1874	-0.0666	0.3328	9.4107	0.6917	16945.88	16984.20	0.9667	0.9552
				(0.0170)		(0.0304)	(0.0153)	(0.1328)	(1.1688)	(0.0311)				
			GAS	0.0397	0.0057	0.0759	-	0.9941	9.3464	-	16961.62	16993.54	0.9862	0.9844
				(0.0204)	(0.0026)	(0.0091)		(0.0021)	(1.1745)					
			FIGAS	0.0417	-	0.0968	-	0.3904	9.5363	0.6993	16965.03	16996.96	0.9894	0.9783
				(0.0202)		(0.0175)		(0.1414)	(1.2135)	(0.0339)				
			GAS+L	0.0355	0.0032	0.0373	0.0593	0.9952	9.4181	-	16942.19	16980.50	0.9508	0.9549
				(0.0020)	(0.0022)	(0.0105)	(0.0128)	(0.0018)	(1.1886)					
			FIGAS+L	0.0468	-	0.0556	0.0881	0.2967	9.6400	0.7029	16941.17	16979.48	0.9522	0.9459
				(0.0182)		(0.0147)	(0.0236)	(0.1589)	(1.2368)	(0.0325)				
GE	0.5358	0.5076	GARCH	0.0025	0.0094	0.0530	-	0.9420	10.2927	-	15132.64	15164.57	1.0000	1.0000
	(0.0284)	(0.0284)		(0.0171)	(0.0033)	(0.0079)		(0.0086)	(1.2672)					
			FIGARCH	0.0057	-	0.3132	-	0.7436	11.4195	0.5444	15158.66	15190.59	1.0363	1.0486
				(0.0170)		(0.0372)		(0.0390)	(1.3823)	(0.0480)				
			EGARCH	-0.0096	0.0035	0.1232	-0.0455	0.9929	10.7233	-	15096.23	15134.54	0.9805	0.9814
				(0.0170)	(0.0020)	(0.0164)	(0.0085)	(0.0021)	(1.4663)					
			FIEGARCH	0.0060	-	0.2163	-0.0740	0.1215	11.9301	0.7078	15077.14	15115.46	0.9767	0.9780
				(0.0145)		(0.0327)	(0.0162)	(0.1568)	(1.7936)	(0.0370)				
			GAS	0.0014	0.0021	0.0651	-	0.9952	10.2605	-	15127.21	15159.13	0.9906	0.9799
				(0.0171)	(0.0019)	(0.0088)		(0.0021)	(1.3151)					
			FIGAS	0.0027	-	0.1108	-	0.0742	11.7506	0.7427	15103.16	15135.09	0.9832	0.9695
				(0.0170)		(0.0174)		(0.1677)	(1.7217)	(0.0395)				
			GAS+L	0.0009	0.0028	0.0429	0.0457	0.9935	10.7694	-	15115.61	15153.93	0.9740	0.9651
				(0.0171)	(0.0020)	(0.0095)	(0.0130)	(0.0023)	(1.4401)					
			FIGAS+L	0.0077	-	0.0752	0.0683	0.1439	12.3729	0.7126	15093.18	15131.49	0.9660	0.9565
				(0.0161)		(0.0165)	(0.0214)	(0.1574)	(1.8834)	(0.0401)				

Note: Standard error between parentheses. $d^{(T)RV}$ denotes the estimate of the fractional difference parameter for the (T)RV measure based on log-periodogram regression method of Geweke and Porter-Hudak (1983). The GPH statistic was estimated at $T^{3/4}$ following Maynard and Phillips (2001). AIC stands for Akaike Information Criterion and SIC stands for Schwarz (or Bayesian) Information Criterion. MAE$^{(T)RV}$ stands for the mean of the absolute errors between the model-based volatility process and (threshold) realized volatility. MAE statistics are related to the GARCH model.

Table 4.B.I: *(Continued)*

Equity	d^{RV}	d^{TRV}	Model	θ	δ	α	γ	β	ν	d	AIC	SIC	MAE^{RV}	MAE^{TRV}
KO	0.5480	0.5271	GARCH	0.0452	0.0038	0.0379	-	0.9603	7.4960	-	13841.91	13873.84	1.0000	1.0000
	(0.0284)	(0.0284)		(0.0150)	(0.0017)	(0.0060)		(0.0064)	(0.7261)					
			FIGARCH	0.0464	-	0.3038	-	0.7848	8.2196	0.5086	13835.42	13867.35	0.9973	0.9969
				(0.0149)		(0.0371)		(0.0376)	(0.7529)	(0.0466)				
			EGARCH	0.0349	0.0018	0.0983	-0.0253	0.9953	7.4173	-	13825.43	13863.74	0.9712	0.9694
				(0.0149)	(0.0012)	(0.0136)	(0.0068)	(0.0018)	(0.7686)					
			FIEGARCH	0.0398	-	0.1510	-0.0383	0.3433	7.6511	0.6958	13817.07	13855.38	0.9784	0.9746
				(0.0139)		(0.0330)	(0.0128)	(0.1766)	(0.7860)	(0.0456)				
			GAS	0.0428	0.0016	0.0544	-	0.9957	7.6291	-	13835.79	13867.72	0.9750	0.9724
				(0.0149)	(0.0011)	(0.0080)		(0.0018)	(0.7911)					
			FIGAS	0.0427	-	0.0726	-	0.4278	7.8584	0.7054	13829.63	13861.56	0.9855	0.9821
				(0.0149)		(0.0195)		(0.1996)	(0.8339)	(0.0519)				
			GAS+L	0.0333	0.0004	0.0206	0.0543	0.9969	8.0062	-	13813.58	13851.89	0.9676	0.9661
				(0.0143)	(0.0009)	(0.0080)	(0.0109)	(0.0014)	(0.8643)					
			FIGAS+L	0.0357	-	0.0363	0.0856	0.3237	8.2752	0.7003	13805.18	13843.49	0.9632	0.9597
				(0.0132)		(0.0129)	(0.0245)	(0.1787)	(0.9216)	(0.0436)				
PG	0.4807	0.5102	GARCH	0.0945	0.0080	0.0499	-	0.9463	7.4017	-	13804.39	13836.32	1.0000	1.0000
	(0.0284)	(0.0284)		(0.0151)	(0.0029)	(0.0080)		(0.0087)	(0.7273)					
			FIGARCH	0.0842	-	0.4001	-	0.7122	8.3683	0.4392	13856.10	13888.03	1.0499	1.0526
				(0.0147)		(0.0311)		(0.0321)	(0.7381)	(0.0294)				
			EGARCH	0.0858	0.0027	0.1255	-0.0290	0.9918	7.5312	-	13790.76	13829.08	0.9610	0.9564
				(0.0175)	(0.0017)	(0.0159)	(0.0089)	(0.0027)	(0.7777)					
			FIEGARCH	0.0902	-	0.1925	-0.0470	0.4065	7.9230	0.6121	13781.02	13819.34	0.9546	0.9477
				(0.0142)		(0.0341)	(0.0146)	(0.1484)	(0.8234)	(0.0442)				
			GAS	0.0958	0.0028	0.0691	-	0.9924	7.7077	-	13795.94	13827.87	0.9621	0.9570
				(0.0151)	(0.0016)	(0.0094)		(0.0027)	(0.7918)					
			FIGAS	0.0959	-	0.1048	-	0.3711	8.1353	0.6368	13790.22	13822.16	0.9599	0.9522
				(0.0150)		(0.0211)		(0.1752)	(0.8705)	(0.0475)				
			GAS+L	0.0900	0.0006	0.0478	0.0458	0.9922	7.7707	-	13787.20	13825.52	0.9533	0.9483
				(0.0150)	(0.0017)	(0.0100)	(0.0145)	(0.0026)	(0.8011)					
			FIGAS+L	0.0890	-	0.0795	0.0649	0.3461	8.2252	0.6319	13780.84	13819.16	0.9396	0.9325
				(0.0149)		(0.0174)	(0.0229)	(0.1595)	(0.8955)	(0.0466)				

It is assumed that $y_t = \theta + \sigma_t z_t$, where z_t is idd variable that has t distribution with ν degrees of freedom as in (4.1). Models for conditional volatility are in their most parsimonious versions: GARCH: $\sigma^2_{t+1} = \delta + \alpha y_t^2 + \beta\sigma_t^2$; FIGARCH: $\sigma^2_{t+1} = \delta + \beta\sigma_t^2 + \left[1 - \beta L - (1-\alpha L)(1-L)^d\right] y^2_{t+1}$; EGARCH: $h_{t+1} = \delta + f(z_t) + \beta h_t$ with $h_t = \log\sigma_t^2$ and with $f(z_t) = \gamma z_t + \alpha[|z_t| - \mathbb{E}|z_t|]$ and $\mathbb{E}|z_t| = \frac{\sqrt{\nu-2}\Gamma((\nu+1)/2)}{(\nu-1)\sqrt{\pi}\Gamma(\nu/2)}$; FIEGARCH: $(1-\beta L)(h_{t+1} - \mu) = (1-L)^{-d} f(z_t)$; GAS: $h_{t+1} = \delta + \alpha\eta_t + \beta h_t$ with η_t defined in (4.8); FIGAS: $(1-\beta L)(h_{t+1} - \mu) = \alpha(1-L)^{-d}\eta_t$; GAS+L: $h_{t+1} = \delta + f(\eta_t) + \beta h_t$ with $f(\eta_t) = \left[\alpha + \gamma\, \mathbf{1}_{\{y_t<0\}}\right]\eta_t$ where $\mathbf{1}_{\{y_t<0\}}$ denotes the indicator function taking value one when $y_t < 0$ and zero else. FIGAS+L: $(1-\beta L)(h_{t+1} - \mu) = (1-L)^{-d} f(\eta_t)$. We do not report estimates of μ as it is not identified when $d \geq 1/2$.

Table 4.B.II: *Maximum likelihood estimates for conditional dependence between daily returns on equities*

Pair	d^{RCorr}	d^{TRCorr}	Model	δ/μ	α	β	d	κ	AIC	SIC	MAE^{RCorr}	MAE^{TRCorr}
AXP/GE	0.5044	0.4974	Fisher-G	0.0074	0.0488	0.9803	-	-	-1220.05	-1200.89	1.0000	1.0000
	(0.0284)	(0.0284)		(0.0029)	(0.0116)	(0.0052)						
			GAS-G	0.0038	0.0356	0.9964	-	-	-1217.33	-1198.17	0.9721	0.9650
				(0.0029)	(0.0081)	(0.0024)						
			FIBase-G	1.0262	0.0306	0.5882	0.5461	-	-1224.87	-1201.21	0.9711	0.9666
				(0.1523)	(0.0432)	(0.5371)	(0.0728)					
			FIGAS-G	0.9501	0.0447	0.4080	0.7138	-	-1221.64	-1196.10	0.9464	0.9413
				(0.1964)	(0.0275)	(0.3840)	(0.0768)					
			Fisher-t	0.0055	0.0446	0.9820	-	9.9418	-1253.73	-1228.19	0.9897	0.9902
				(0.0026)	(0.0110)	(0.0050)		(1.8900)				
			GAS-t	0.0032	0.0377	0.9969	-	9.8904	-1248.64	-1223.09	0.9553	0.9508
				(0.0029)	(0.0094)	(0.0025)		(1.9109)				
			FIBase-t	1.0596	0.0367	0.5558	0.5856	8.4947	-1231.26	-1199.33	0.9507	0.9456
				(0.1448)	(0.0425)	(0.5597)	(0.0725)	(1.2691)				
			FIGAS-t	0.9548	0.0738	0.2514	0.7109	9.9692	-1252.60	-1220.67	0.9371	0.9345
				(0.2091)	(0.0297)	(0.3337)	(0.0741)	(1.9385)				
AXP/KO	0.4455	0.4407	Fisher-G	0.0053	0.0446	0.9814	-	-	-465.66	-446.51	1.0000	1.0000
	(0.0284)	(0.0284)		(0.0022)	(0.0105)	(0.0048)						
			GAS-G	0.0033	0.0260	0.9951	-	-	-449.43	-430.27	0.9751	0.9642
				(0.0020)	(0.0066)	(0.0027)						
			FIBase-G	1.3636	0.0779	0.9978	-0.1548	-	-456.99	-431.45	0.9847	0.9847
				(0.4436)	(0.0381)	(0.0019)	(0.1322)					
			FIGAS-G	0.7444	0.0291	0.9970	-0.0533	-	-448.77	-423.22	0.9982	0.9881
				(0.1036)	(0.0182)	(0.0035)	(0.1836)					
			Fisher-t	0.0044	0.0345	0.9840	-	8.0571	-501.56	-476.02	0.9786	0.9776
				(0.0021)	(0.0091)	(0.0046)		(1.4356)				
			GAS-t	0.0028	0.0273	0.9957	-	7.4992	-494.06	-468.52	0.9562	0.9450
				(0.0020)	(0.0078)	(0.0027)		(1.2479)				
			FIBase-t	0.8951	0.0446	0.9975	-0.2496	8.7724	-489.95	-458.02	1.0834	1.0753
				(0.1362)	(0.0289)	(0.0023)	(0.1696)	(1.6325)				
			FIGAS-t	0.7264	0.0247	0.9953	0.0185	7.4793	-493.07	-461.14	0.9788	0.9686
				(0.1233)	(0.0174)	(0.0049)	(0.2127)	(1.2314)				

Note: Standard error between parentheses. $d^{(T)RCorr}$ denotes the estimate of the fractional difference parameter for the (T)RCorr measure based on log-periodogram regression method of Geweke and Porter-Hudak (1983). AIC stands for Akaike Information Criterion and SIC stands for Schwarz (or Bayesian) Information Criterion. MAE$^{(T)RCorr}$ stands for the mean of the absolute errors between the model-based correlation process and (threshold) realized correlation. Suffixes G and t stand for Gaussian and t copula respectively. The Fisher dynamics for correlation has the form as in (4.28); the FIBase model is given in (4.29). Details of GAS and FIGAS models are discussed in Section 4.2.4.

Table 4.B.II: *(Continued)*

Pair	d^{RCorr}	d^{TRCorr}	Model	δ/μ	α	β	d	κ	AIC	SIC	MAE^{RCorr}	MAE^{TRCorr}
AXP/PG	0.3931	0.3805	Fisher-G	0.0019	0.0222	0.9919	-	-	-496.68	-477.52	1.0000	1.0000
	(0.0284)	(0.0284)		(0.0010)	(0.0052)	(0.0021)						
			GAS-G	0.0759	0.0894	0.9037	-	-	-481.05	-461.89	1.0769	1.0648
				(0.0261)	(0.0193)	(0.0322)						
			FIBase-G	1.0220	0.1785	0.5490	0.2380	-	-472.28	-446.74	1.0258	1.0184
				(0.1210)	(0.0868)	(0.2542)	(0.1074)					
			FIGAS-G	0.7700	0.0713	0.6175	0.4557	-	-492.29	-466.75	1.0177	1.0073
				(0.0927)	(0.0201)	(0.1566)	(0.1171)					
			Fisher-t	0.0019	0.0183	0.9923	-	9.3959	-526.31	-500.77	0.9897	0.9886
				(0.0011)	(0.0050)	(0.0023)		(1.8759)				
			GAS-t	0.0615	0.0897	0.9208	-	8.8568	-516.86	-491.32	1.0504	1.0373
				(0.0282)	(0.0238)	(0.0356)		(1.6507)				
			FIBase-t	0.8474	0.1211	0.4293	0.3012	8.6891	-507.94	-475.97	1.0913	1.0820
				(0.0814)	(0.0707)	(0.3497)	(0.1167)	(1.5909)				
			FIGAS-t	0.7536	0.0679	0.5654	0.5285	8.9630	-527.97	-496.04	0.9923	0.9807
				(0.1188)	(0.0250)	(0.2178)	(0.1402)	(1.6783)				
GE/KO	0.4403	0.4623	Fisher-G	0.0050	0.0343	0.9853	-	-	-753.75	-734.59	1.0000	1.0000
	(0.0284)	(0.0284)		(0.0023)	(0.0086)	(0.0042)						
			GAS-G	0.0031	0.0227	0.9964	-	-	-755.65	-736.49	0.9981	0.9907
				(0.0019)	(0.0053)	(0.0019)						
			FIBase-G	1.2045	0.2136	0.2200	0.3179	-	-722.27	-696.79	1.0058	1.0041
				(0.1165)	(0.1211)	(0.3983)	(0.0965)					
			FIGAS-G	0.8169	0.0573	0.3507	0.6526	-	-747.06	-721.51	0.9981	0.9895
				(0.1399)	(0.0274)	(0.3559)	(0.0852)					
			Fisher-t	0.0051	0.0311	0.9852	-	8.1738	-799.50	-773.96	1.0001	1.0005
				(0.0025)	(0.0086)	(0.0046)		(1.3605)				
			GAS-t	0.0035	0.0271	0.9959	-	8.1836	-802.58	-777.04	0.9928	0.9851
				(0.0025)	(0.0076)	(0.0026)		(1.3525)				
			FIBase-t	0.9471	0.2029	0.0551	0.2983	7.9604	-772.76	-740.83	1.0722	1.0664
				(0.0677)	(0.0972)	(0.4744)	(0.0902)	(1.2430)				
			FIGAS-t	0.8045	0.0617	0.4051	0.6524	8.0842	-797.67	-765.74	0.9925	0.9841
				(0.1635)	(0.0328)	(0.3799)	(0.0976)	(1.3093)				

Table 4.B.II: *(Continued)*

Pair	d^{RCorr}	d^{TRCorr}	Model	δ/μ	α	β	d	κ	AIC	SIC	MAE^{RCorr}	MAE^{TRCorr}
GE/PG	0.4461	0.4343	Fisher-G	0.0061	0.0398	0.9823	-	-	-718.40	-699.24	1.0000	1.0000
	(0.0284)	(0.0284)		(0.0030)	(0.0103)	(0.0054)						
			GAS-G	0.0045	0.0227	0.9950	-	-	-712.12	-692.96	1.0049	0.9968
				(0.0024)	(0.0057)	(0.0026)						
			FIBase-G	1.3024	0.1721	0.2647	0.3847	-	-694.18	-668.64	0.9954	0.9937
				(0.1463)	(0.1997)	(0.8444)	(0.0906)					
			FIGAS-G	0.9033	0.0983	0.3971	0.4602	-	-714.14	-688.59	1.0289	1.0208
				(0.0872)	(0.0274)	(0.2106)	(0.0884)					
			Fisher-t	0.0073	0.0332	0.9820	-	6.8425	-787.20	-761.65	1.0128	1.0128
				(0.0041)	(0.0110)	(0.0072)		(0.9548)				
			GAS-t	0.0048	0.0266	0.9948	-	6.7104	-787.18	-761.64	1.0150	1.0085
				(0.0032)	(0.0081)	(0.0034)		(0.9055)				
			FIBase-t	0.9489	0.1655	0.0380	0.3105	6.6049	-771.99	-740.04	1.0651	1.0588
				(0.0693)	(0.0916)	(0.4948)	(0.0831)	(0.8665)				
			FIGAS-t	0.9175	0.1420	0.2642	0.4609	6.6904	-793.50	-761.57	1.0423	1.0349
				(0.1018)	(0.0374)	(0.2356)	(0.0909)	(0.8855)				
KO/PG	0.3937	0.4060	Fisher-G	0.0109	0.0412	0.9798	-	-	-897.53	-878.37	1.0000	1.0000
	(0.0284)	(0.0284)		(0.0061)	(0.0137)	(0.0085)						
			GAS-G	0.0237	0.0499	0.9780	-	-	-906.37	-887.21	1.0042	1.0005
				(0.0093)	(0.0110)	(0.0085)						
			FIBase-G	1.4021	0.1501	0.4085	0.2910	-	-872.77	-847.23	0.9931	0.9927
				(0.1260)	(0.1287)	(0.5020)	(0.0821)					
			FIGAS-G	1.0901	0.0434	0.9699	0.0779	-	-905.64	-880.10	1.0091	1.0056
				(0.0648)	(0.0183)	(0.0293)	(0.2317)					
			Fisher-t	0.0138	0.0479	0.9739	-	7.1059	-957.83	-932.29	0.9934	0.9935
				(0.0075)	(0.0149)	(0.0103)		(1.0506)				
			GAS-t	0.0305	0.0659	0.9715	-	7.5000	-961.99	-936.45	0.9959	0.9920
				(0.0122)	(0.0147)	(0.0111)		(1.1656)				
			FIBase-t	1.2367	0.0444	0.9967	0.0554	7.6501	-955.85	-923.88	1.0805	1.0793
				(0.1046)	(0.0232)	(0.0021)	(0.0961)	(1.1422)				
			FIGAS-t	1.0744	0.0519	0.9449	0.1727	7.4838	-961.45	-929.52	0.9943	0.9904
				(0.0793)	(0.0181)	(0.0447)	(0.2067)	(1.1564)				

Chapter 5

Modeling daily covariance: A joint framework for low and high-frequency based measures

Abstract: This chapter develops a novel approach for the joint modeling of daily financial returns and daily realized covariance. We make distributional assumptions about returns and realized covariance where both distributions depend on the latent covariance matrix. The key novelty is the dynamic specification for modeling covariance matrix based on a combination of past realized measures of daily covariance matrices and past outer-products of daily returns. Consequently, the model for daily covariance makes use of both low and high-frequency data. We show that in our model specification each individual innovation driving the covariance matrix exploits full likelihood information. Each model formulation allows for cross-asset effects with the consequence that the number of parameters can be kept low in high dimensions. We apply our novel approach to a NYSE equity dataset, and we show that our model outperforms some alternative frameworks.

5.1 Introduction

The estimation of covariance matrices of financial assets has been traditionally based on daily financial returns using multivariate versions of the well-known generalized autoregressive conditional heteroskedasticity (GARCH) models or stochastic volatility (SV) models. Both models consider time-varying second moment of the return data, where the updates for covariance parameters are specified as autoregressive moving average processes which imply persistence properties for variances and covariances. The GARCH and SV models differ in the way the innovations to time-varying covariance parameters are defined. While

in the SV models covariance parameters are driven by stochastic innovations, then in the GARCH models dynamics of covariance parameters is determined by the outer-product of lagged returns. A wide range of both multivariate GARCH and SV models have been developed and applied extensively to empirical data in the recent years. For a comprehensive overview of multivariate GARCH models, we refer to Bauwens, Laurent, and Rombouts (2006) and Silvennoinen and Teräsvirta (2009). A review of multivariate SV models is provided by Asai, McAleer, and Yu (2006).

The main shortcoming of traditional GARCH and SV models is that they make use of price observations recorded at daily intervals to infer about the current level of volatilities and covariances of assets. Given the increasing availability of high-frequency intraday data, the use of only low-frequency daily data appears to be suboptimal. One important consequence is that models based on daily data do not adapt quickly enough to changes in volatilities which is key to track the financial risk in a timely manner; see Andersen, Bollerslev, Diebold, and Labys (2003) for more discussion. For this reason, various attempts have been made to integrate the high-frequency intraday data into the measuring and modeling financial volatility. For instance, the information from high-frequency data can be incorporated by adding it in the form of an explanatory variable to the GARCH or SV volatility dynamics; see Engle (2002b) and Koopman, Jungbacker, and Hol (2005) for GARCH and SV approaches respectively.

With the advent of high-frequency data, one can estimate ex-post daily variation of a single asset price changes with the so-called realized volatility measures; see Andersen and Bollerslev (1998a), Andersen, Bollerslev, Diebold, and Labys (2001) and Barndorff-Nielsen and Shephard (2002). Inherent to high-frequency data is the microstructure noise (bid-ask bounce, price discreteness etc.) which leads to bias. A number of related quantities have been developed to restore consistency; see Aït-Sahalia, Mykland, and Zhang (2005), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008), Jacod, Li, Mykland, Podolskij, and Vetter (2009b), and references therein. In the case of multiple assets, the high-frequency realized measures of covariance of assets have been also proposed; see Malliavin and Mancino (2002), Hayashi and Yoshida (2005), Christensen, Kinnebrock, and Podolskij (2010), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011), Griffin and Oomen (2011), and references therein.

Computation of realized measures from high-frequency data permits the use of more standard time series techniques for modeling and forecasting financial volatility. Andersen et al. (2001) have explored the use of the autoregressive model to analyze the time series of realized volatilities. They have found considerable improvements in the volatility forecasts over more standard GARCH models. Recently, some new promising models have been proposed that rely on time series of realized measures. Gourieroux, Jasiak, and Sufana (2009) have proposed the (non-central) Wishart autoregressive model for realized covariance matrix. Asai and So (2010) and Golosnoy, Gribisch, and Liesenfeld (2010) have

proposed alternative dynamic formulation for covariance parameters with the underlying Wishart density. Chiriac and Voev (2011) and Bauer and Vorkink (2011) have proposed models for the realized covariances using appropriate transformations to ensure positive definiteness of the covariance matrix. In this chapter, we rely also on the Wishart density, while the novelty concerns our formulation for covariance which combines information from daily as well as intradaily price observations.

An approach that combines possibly several measures of volatility based on low- and high-frequency data is developed by Engle and Gallo (2006). They model jointly close-to-close returns, range and realized variance with their multiplicative error model (MEM) where each measure has its own dynamics for the update of volatility augmented with lagged values of other two measures. They show that the combination of these three measures of volatility can bring gains when making medium-run volatility forecasts. Shephard and Sheppard (2010) propose a model for close-to-close returns and realized kernel with GARCH specification for two latent variance processes and show that their model quickly adapts to changes in the level of the volatility process. Hansen, Huang, and Shek (2011) propose the realized GARCH model where daily returns and ex-post measure of daily variation are both linked to the latent volatility. They classify their model as complete since variation in returns and realized volatility are jointly modeled.

This chapter contributes to the recent developments in the joint modeling of returns and realized measures. We consider modeling of the covariance matrix of financial assets. Our aim is to provide dynamics for the latent daily covariance matrix by taking advantage of both low and high-frequency data. To highlight the idea more in detail, we propose a specification for the latent covariance matrix as a function of past high-frequency realized measures of daily covariance matrices and past outer-products of daily vector returns. Both realized measure and outer-product of daily returns contain a noisy signal about the current covariance parameters. We adopt the generalized autoregressive score (GAS) framework of Creal, Koopman, and Lucas (2010) to model jointly daily returns and the multivariate realized kernel of asset prices proposed by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011). Within the GAS modeling framework, we need to impose distributional assumptions for the vector of daily returns as well as for the multivariate realized kernel. In this chapter we assume a multivariate normal density for returns and a Wishart density for the multivariate realized kernel. Both distributions depend on a latent covariance matrix. The resulting novel model relies on a dynamic specification for the covariance matrix that is driven by the scaled score of the log-likelihood function. The score is shown to be a combination of the outer-product of daily returns and the realized kernel. The realized measure receives most weight by construction, and this weight is associated with the number of degrees of freedom in the Wishart distribution. Indeed, realized kernel is a more accurate measure of the latent covariance given that it exploits intraday high-frequency data of assets.

The key distinguishing feature of our model specification is that each innovation driving the covariance matrix exploits full likelihood information. We demonstrate that even a simplified scalar specification allows for a complex interdependence between all variances and covariances such that each one-step ahead prediction is driven by own as well as cross-asset effects. This feature distinguishes our model from a low frequency GARCH model driven by outer-product of daily returns, but also from the high-frequency HEAVY model driven by the realized covariance. Both models require full parametrization to allow for cross-asset effects. Our model is parsimonious which is a convenient property for multivariate volatility models.

The remainder of this chapter is organized as follows. In Section 2 we review high-frequency based measurement of covariance matrix of assets. Section 3 sets out the notation and assumptions. In this section we introduce the score-based model for daily covariance. In Section 4, we conduct simulations to study to the likelihood-based estimation. Section 5 presents the results of our empirical analysis for a set of NYSE equities, while Section 6 concludes this chapter. The proofs of the main results in this chapter are given in the Appendix.

5.2 Measuring Covariation

We study a k-dimensional log price process $Y(\tau) = \big(Y^1(\tau), Y^2(\tau), \ldots, Y^k(\tau)\big)$, where $Y^i(\tau)$ denotes time τ logarithmic price of asset i, and $\tau \in [0,1]$. The price observations are assumed to be observed at discrete times that for asset i are denoted as $\tau_0^{(i)}, \tau_1^{(i)}, \ldots$. Hence, the price observations of asset i at the observation times are given by $Y^i(\tau_j^i)$ for $j = 0, 1, \ldots, N^i(1)$, where $N^i(\tau)$ counts the number of price observations for the ith asset up to time τ. The observed price observations Y are assumed be driven by the latent efficient prices X and microstructure noise U. The efficient price process X is further assumed to be given by the Brownian semimartingale,

$$X_\tau = \int_0^\tau \mu_s \mathrm{d}s + \int_0^\tau \sigma_s \mathrm{d}W_s, \tag{5.1}$$

where μ_τ is a vector of elements which are predictable locally bounded drifts, σ_τ is a cadlag volatility matrix process and W_τ is a vector of independent Brownian motions.

The ex-post covariation over the $[0,1]$ interval is defined as

$$[X]_1 = \int_0^1 \Sigma_\tau \, \mathrm{d}\tau, \qquad \text{with} \qquad \Sigma_\tau = \sigma_\tau \sigma_\tau'. \tag{5.2}$$

where

$$[X]_1 = \operatorname*{plim}_{n\to\infty} \sum_{j=1}^{n} (X_{\tau_j} - X_{\tau_{j-1}}) \, (X_{\tau_j} - X_{\tau_{j-1}})', \tag{5.3}$$

is quadratic variation for any sequence of synchronized partitions $0 = \tau_0 < \ldots < \tau_n = 1$ with $\tau_j - \tau_{j-1} \to 0$ or $n \to \infty$.

5.2.1 Covariance estimation in a frictionless market

First, for simplicity, we consider the case where there is no microstructure noise implying $Y = X$. The primary concern in covariation estimation is synchronization of price observations. Here, we focus on the standard previous-tick method which ensures that returns are measured over a common time grid in calendar time; see Hansen and Lunde (2006). In this method, each sampling point for each asset is determined by the most recently observed price. In this way, we can easily construct n synchronized vector returns $y_{\tau_j} = Y_{\tau_j} - Y_{\tau_{j-1}}$ for $j = 1, \ldots, n$. The covariation (5.2) can be estimated consistently by the realized covariance given by

$$RC_1(Y) = \sum_{j=1}^{n} y_{\tau_j} y'_{\tau_j}. \tag{5.4}$$

The use of the previous-tick method facilitates synchronization of the data, but it does not eliminate the non-trading problem. The non-synchronous trading has been found to bias the covariance measures towards zero, and this phenomenon is often referred to as the Epps effect; see Epps (1979). Consequently, the realized covariances computed from returns measured over short time periods underestimate the degree of dependence between assets; see Griffin and Oomen (2011) for more discussion.

To alleviate the possible underestimation of covariance, we make use of subsampling. The idea of subsampling for construction of realized volatilities was suggested by Aït-Sahalia et al. (2005). The motivation behind subsampling is that we do not rather sample too frequently to mitigate noise and to avoid non-trading bias, but we still want to use all data. In principle, we can jitter the initial point q-second adjacent intervals over n observations per day in q different ways resulting in q series of returns

$$Y_{\tau_q} - Y_{\tau_0}, Y_{\tau_{2q}} - Y_{\tau_q}, \ldots,$$
$$Y_{\tau_{q+1}} - Y_{\tau_1}, Y_{\tau_{2q+1}} - Y_{\tau_{q+1}}, \ldots,$$

and so on. The resulting subsampling estimator is defined as the average of q realized covariances obtained as in (5.4). We will use the subsampled realized covariance to evaluate our model in the empirical illustration.

5.2.2 Covariance estimation robust to microstructure noise

To estimate the covariation consistently in the presence of microstructure noise implying $Y \neq X$, we apply the multivariate realized kernel proposed by Barndorff-Nielsen et al.

(2011), which is the first estimator to simultaneously address microstructure noise and non-synchronous trading while guaranteeing consistency and positive semi-definiteness. Market microstructure effects are modeled through an additive noise component as

$$Y^i(\tau_j^i) = X^i(\tau_j^i) + U^i(\tau_j^i), \tag{5.5}$$

where $U^i(\tau_j^i)$ is covariance stationary and satisfies $\mathbb{E}[U^i(\tau_j^i)] = 0$. The market microstructure effects can be time dependent and correlated with efficients innovations, see Barndorff-Nielsen et al. (2011) for more discussion.

To estimate the covariation, one needs to regularize the observation times. Barndorff-Nielsen et al. (2011) propose to synchronize the data by the so-called refresh time sampling (rts). Once all assets have been traded, the most recent prices are used to form the refresh timescale. In particular, the first refresh time is given by

$$\tau_1^{rts} = \max\big(\tau_1^1, \tau_1^2, \ldots, \tau_1^k\big), \tag{5.6}$$

and the subsequent refresh times are given by

$$\tau_{j+1}^{rts} = \max\big(\tau^1_{N^1(\tau_j^{rts})+1}, \tau^2_{N^2(\tau_j^{rts})+1}, \ldots, \tau^k_{N^k(\tau_j^{rts})+1}\big). \tag{5.7}$$

The asymptotic theory of kernel estimator dictates that we need to average m prices at the very beginning and end of the day to obtain a consistent estimator. We follow the suggestion of Barndorff-Nielsen et al. (2011) and set $m = 2$ implying that the first (last) price is the average of the first (last) two rts prices. The resulting number of rts price observations is denoted by n^{rts}. To gauge the potential loss of the data due to refresh time sampling, we follow Barndorff-Nielsen et al. (2011) by computing statistics

$$p = \frac{kn^{rts}}{\sum_{i=1}^k N^i(1)}, \tag{5.8}$$

which measures the ratio of the size of the retained data due to synchronization over the original size of the database. In principle, the loss of the data is larger, the less liquid the assets are. In our empirical application we provide statistics (5.8) for $k \in \{2, 5, 15\}$. Having synchronized high-frequency price observations and having dealt with end effects, we compute a vector of returns $y_{\tau_j^{rts}} = Y_{\tau_j^{rts}} - Y_{\tau_{j-1}^{rts}}$ for $j = 1, \ldots, n^{rts}$. We next construct the positive semi-definite multivariate realized kernel defined as

$$RK_1(Y) = \sum_{h=-H}^{H} g\left(\frac{h}{H+1}\right)\Omega_h, \tag{5.9}$$

where the weight function g is given by the Parzen kernel function,

$$g(x) = \begin{cases} 1 - 6x^2 + 6x^3 & \text{for} \quad 0 \le x \le 1/2 \\ 2(1-x)^3 & \text{for} \quad 1/2 \le x \le 1 \\ 0 & \text{for} \quad x > 1 \end{cases}, \tag{5.10}$$

and where Ω_h is the hth realized autocovariance matrix,

$$\Omega_h = \sum_{j=h+1}^{n^{rts}} y_{\tau_j^{rts}} y'_{\tau_{j-h}^{rts}}, \qquad \text{for} \quad h \geq 0, \tag{5.11}$$

and $\Omega_h = \Omega'_h$ for $h < 0$, see Barndorff-Nielsen et al. (2011). The choice of the single bandwidth parameter H is here based on the average of individual bandwidths given by $H^i = c\xi_i^{4/5}(n^{rts})^{3/5}$, where $c = 3.5134$, $\xi_i = \varpi_i^2/\sqrt{IQ_i}$ denotes the noise-to-signal ratio, ϖ_i^2 is microstructure noise variance, and IQ_i is the integrated quarticity measure. In our empirical implementation, we follow Barndorff-Nielsen et al. (2009) and estimate

$$\widehat{\varpi}_i^2 = \frac{RV^i_{dense}}{2\bar{N}^i(1)} \qquad \text{and} \qquad IQ_i = \left(RV^i_{sparse}\right)^2, \tag{5.12}$$

where $\bar{N}^i(1)$ counts the number of non-zero returns used to compute RV^i_{dense} with RV^i_{dense} denoting the realized variance estimated with data sampled every qth trade; here $q = 25$ which suffices to mitigate noise effects; RV^i_{sparse} denotes the subsampled realized variance based on 20 minute returns. Since $q = 25$, then by jittering the starting point we obtain 25 distinct measures of variance of microstructure noise which we average to get the subsampled version. The integrated quarticity is here approximated by the square of integrated variance whose estimation is more accurate. The global bandwidth is simply set to $H = k^{-1}\sum_{i=1}^{k} H^i$; for some other possible choices, see Barndorff-Nielsen et al. (2011).

5.3 Score-based Modeling

Having computed the series of daily realized kernels, we turn our attention to modeling the covariance matrix of a set of financial assets. Our modeling approach focuses on the joint modeling of daily returns and daily realized covariances with the aim to develop a model formulation for covariance matrix which exploits both low and high-frequency data. We start with matrix notations and modeling assumptions. We then describe the details of modeling strategy and we provide details of our model.

5.3.1 Matrix notation and definitions

To develop our results in this chapter, we adopt the following matrix notation and definitions. For convenience, we restrict general results to square matrices as these are used in this article. If A is a $k \times k$ matrix and B is a $k \times k$ matrix, then the Kronecker product denoted by $A \otimes B$ is the $k^2 \times k^2$ block matrix. The $vec(A)$ operator stacks the columns of matrix A into $k^2 \times 1$ column vector, while $vech(A)$ stacks the the lower triangular part including diagonal into $k^* \times 1$ column vector, with $k^* = k(k+1)/2$. The unit matrix of

size k is denoted as I_k. The commutation matrix K_k is the $k^2 \times k^2$ matrix which, for any $k \times k$ matrix A, transforms $vec(A)$ into $vec(A')$ such that

$$K_k \, vec(A) = vec(A'). \tag{5.13}$$

The duplication matrix D_k of dimension $k^2 \times k^*$ and the elimination matrix L_k of dimension $k^* \times k^2$ are defined as

$$D_k \, vech(A) = vec(A) \qquad \text{and} \qquad L_k \, vec(A) = vech(A) \tag{5.14}$$

for a symmetric $k \times k$ matrix A and with $L_k = (D_k' D_k)^{-1} D_k'$ being the Moore-Penrose inverse of the duplication matrix D_k. If a $k \times k$ matrix B is lower triangular, then duplication matrix $\mathscr{D}_k$ of dimension $k^2 \times k^*$ and the elimination matrix $\mathscr{L}_k$ of dimension $k^* \times k^2$ are defined as

$$\mathscr{D}_k \, vech(B) = vec(B) \qquad \text{and} \qquad \mathscr{L}_k \, vec(B) = vech(B), \tag{5.15}$$

with $\mathscr{D}_k' \mathscr{D}_k = I_{k^*}$ and $\mathscr{L}_k = \mathscr{D}_k'$. The difference between L_k and $\mathscr{L}_k$ arises because $vec(B)$ contains zeros; more results and properties can be found in Magnus and Neudecker (1988) and Seber (2007).

5.3.2 Modeling assumptions

Let $r_t \in \mathbb{R}^k$ denote vector of daily (demeaned) returns for k assets and let the $RK_t \in \mathbb{R}^{k \times k}$ denote the covariance matrix of k assets on day t, with $t = 1, \ldots, T$. Let $\mathcal{F}_{t-1}$ be the sigma field generated by the past values of daily returns and covariances. We assume

$$r_t | \mathcal{F}_{t-1} \sim \mathcal{N}_k(0, P_t), \tag{5.16}$$

$$RK_t | \mathcal{F}_{t-1} \sim \mathcal{W}_k(\nu, V_t/\nu), \tag{5.17}$$

where $\mathcal{N}_k$ denotes k-variate normal distribution here with mean zero and symmetric positive definite covariance matrix P_t, and where $\mathcal{W}_k$ denotes the k-variate standard Wishart distribution with symmetric positive definite scale matrix V_t and where ν is the number of degrees of freedom, with $\nu \geq k$. We further assume that P_t and V_t are both measurable with respect to $\mathcal{F}_{t-1}$. Note that distributional assumption (5.16) implies that outer product of daily returns is distributed as

$$r_t r_t' \sim \mathcal{W}_k(1, P_t), \tag{5.18}$$

which is singular Wishart distribution with 1 degree of freedom, see Srivastava (2003), as by construction matrix $r_t r_t'$ has rank one. If r_t is the vector of close-to-close returns, then P_t measures overnight variation along with intraday variation, while V_t measures the

intraday variation. If r_t is the vector of open-to-close returns, then both P_t and V_t measure the variation over a particular trading day.[1]

The measurement equations are given by

$$r_t = P_t^{1/2}\varepsilon_t, \tag{5.19}$$
$$RK_t = V_t^{1/2}\eta_t V_t^{1/2}, \tag{5.20}$$

where $X_t^{1/2}$ denotes the square root matrix of X_t; the innovations are assumed to be *iid* distributed as

$$\varepsilon_t \sim \mathcal{N}_k(0, I_k), \tag{5.21}$$
$$\eta_t \sim \mathcal{W}_k(\nu, I_k/\nu). \tag{5.22}$$

Note that ε_t is vector-valued while η_t is matrix-valued random variable. It is natural to assume that variation of conditional variance of realized returns and of conditional mean of realized covariance share the same source. For this reason, we impose the following structure,

$$P_t = \Lambda^{1/2} V_t \Lambda^{1/2}, \tag{5.23}$$

where we assume $\Lambda^{1/2}$ is $k \times k$ diagonal matrix with $\lambda_{ii} = \lambda_i > 0$, $i = 1, \ldots, k$. In principle, we could propose two separate models for P_t and V_t and link them through the lagged values of one each other in a similar way as in the ARMA like models with additional exogenous variables. This modeling strategy is used by Noureldin et al. (2011) who proposed the HEAVY model. This approach would imply explicit modeling of two latent matrix-variate variables P_t and V_t. Instead, we aim to provide a single updating equation for V_t which will contain all relevant information for making predictions, while P_t is simply related to V_t. We expect that when r_t is a vector of close-to-close returns the individual elements of $\Lambda^{1/2}$ should be larger than one which allows to capture overnight variation.

Our set of distributional assumptions implies the following,

$$\mathbb{E}[r_t r_t' | \mathcal{F}_{t-1}] = \Lambda^{1/2} V_t \Lambda^{1/2} \tag{5.24}$$
$$\mathbb{E}[RK_t | \mathcal{F}_{t-1}] = V_t, \tag{5.25}$$

and

$$\mathbb{V}ar[vec(r_t r_t') | \mathcal{F}_{t-1}] = (I_{k^2} + K_k)\big(\Lambda^{1/2} \otimes \Lambda^{1/2}\big)\big(V_t \otimes V_t\big)\big(\Lambda^{1/2} \otimes \Lambda^{1/2}\big), \tag{5.26}$$
$$\mathbb{V}ar[vec(\Lambda^{-1/2} r_t r_t' \Lambda^{-1/2}) | \mathcal{F}_{t-1}] = (I_{k^2} + K_k)\big(V_t \otimes V_t\big), \tag{5.27}$$
$$\mathbb{V}ar[vec(RK_t) | \mathcal{F}_{t-1}] = \nu^{-1}(I_{k^2} + K_k)\big(V_t \otimes V_t\big). \tag{5.28}$$

[1] It is rather standard to exclude the overnight return when computing realized measures, while modeling of daily returns is based on both open-to-close and close-to-close returns.

The results of (5.26) and (5.28) follow directly from Magnus and Neudecker (1979). Note that the result in (5.24) corresponds to conditional second moment, while the results in (5.26) and (5.27) correspond to conditional fourth moment (kurtosis) of returns, possibly adapted for overnight variation. It is also worth noting that conditional second moment of realized covariance (5.28) provides model-implied volatilities-of-volatilities and volatility spillover effects.

5.3.3 Modeling strategy

The two measurement equations (5.19) and (5.20) originate from two different distributions, however they both depend on the same unobserved covariance matrix. We denote the observation densities of daily returns and daily covariances as given by

$$r_t \sim p_r(r_t|f_t, \mathcal{F}_{t-1}; \psi), \tag{5.29}$$
$$RK_t \sim p_{RK}(RK_t|f_t, \mathcal{F}_{t-1}; \psi), \tag{5.30}$$

for $t = 1, \ldots, T$, and where f_t is a vector of latent time-varying parameters linked to covariance matrix at time t, $\mathcal{F}_t$ is the information set containing all observations (both returns and realized covariances) up to time t, and ψ is a vector of static unknown parameters. In particular, our observations densities are given by (5.16) and (5.17) for returns and realized covariances, respectively, and both distributions are assumed to be driven by a common covariance matrix V_t. The vector f_t is of dimension k^* and contains unique elements of Cholesky factors of covariance matrix V_t. We assume that innovations to returns and realized covariances are independent conditional on f_t and on the information set $\mathcal{F}_{t-1}$. The log-likelihood then reads

$$\mathcal{L} = \sum_{t=1}^{T} \Big(\log p_r(r_t|f_t, \mathcal{F}_{t-1}; \psi) + \log p_{RK}(RK_t|f_t, \mathcal{F}_{t-1}; \psi) \Big). \tag{5.31}$$

Our assumption rules out the possibility of leverage effect in the sense that innovations to returns and volatilities are not allowed to be negatively correlated at time t, i.e. $\mathbb{E}[\varepsilon_{it}\eta_{jt}] = 0$, for any i, j in (5.19) and (5.20). However, any asymmetric effects can be introduced when specifying updating equation for latent f_t as we discuss below.

The dynamics of f_t is modeled as the vector autoregressive moving average process,

$$f_{t+1} = \omega + \sum_{i=1}^{p} B_i f_{t-i+1} + \sum_{j=1}^{q} A_j s_{t-j+1}, \tag{5.32}$$

where s_t is a mean-zero and finite variance martingale difference sequence, ω is an $k^* \times 1$ vector of constants, B_j and A_i are $k^* \times k^*$ matrices with coefficients. The parameters ω, $B_1, \ldots, B_p$, $A_1, \ldots, A_q$ and some possible density specific unknown parameters are all

collected in the static parameter vector ψ. The initial value f_0 is fixed at the unconditional mean of the stationary process f_t. The specification (5.32) can be extended to incorporate some exogenous variables and/or other functions of lagged endogenous variables.

Our modeling approach follows Creal et al. (2010) by setting the innovation s_t equal to the scaled score of the local log-likelihood with respect to the parameters of interest f_t. In particular, the score vector takes an additive form given by

$$\nabla_t = \frac{\partial \log p_r(r_t|f_t, \mathcal{F}_{t-1}; \psi)}{\partial f_t} + \frac{\partial \log p_{RK}(RK_t|f_t, \mathcal{F}_{t-1}; \psi)}{\partial f_t} = \nabla_{r,t} + \nabla_{RK,t}, \tag{5.33}$$

which corresponds to the sum of individual scores. The scaling term is based on the (function of) expected information matrix given similarly by an additive form,

$$\mathcal{I}_t = \mathbb{E}[\nabla_{r,t}\nabla'_{r,t}|\mathcal{F}_{t-1}] + \mathbb{E}[\nabla_{RK,t}\nabla'_{RK,t}|\mathcal{F}_{t-1}]. \tag{5.34}$$

The innovation term is now defined as

$$s_t = \mathcal{I}_t^{-1/2}\nabla_t, \tag{5.35}$$

such that $\mathbb{E}[s_t|\mathcal{F}_{t-1}] = 0$ and $\mathbb{E}[s_t s'_t|\mathcal{F}_{t-1}] = I_{k^*}$. In this approach, the one-step ahead prediction of latent parameters f_t is based on the scaled score that exploits the full likelihood function at time t. Equations (5.29), (5.30), (5.32) and (5.35) formulate the Generalized Autoregressive Score model for the time-varying parameter vector f_t with orders p and q abbreviated as GAS(p, q). We select the square root of the inverse of expected information matrix to scale the score; for other possible choices see Creal et al. (2010). This choice implies that variance of the innovation term s_t is time invariant, but importantly it appears to provide the most stable model formulation in terms of model estimation.

5.3.4 The main result

Given the measurement equations (5.19) and (5.20) for realized return vector and realized covariance matrix respectively, we can formulate the GAS-based specification for unobserved covariance matrix V_t. The log-likelihood at time t reads

$$\mathcal{L}_t = \mathcal{L}_{1,t} + \mathcal{L}_{2,t}, \tag{5.36}$$

where

$$\mathcal{L}_{1,t} = d(k) - \frac{1}{2}\log|\Lambda^{1/2}V_t\Lambda^{1/2}| - \frac{1}{2}tr\left((\Lambda^{1/2}V_t\Lambda^{1/2})^{-1}r_t r'_t\right) \tag{5.37}$$

$$\mathcal{L}_{2,t} = d(k,\nu) + \frac{\nu-k-1}{2}\log|RK_t| - \frac{\nu}{2}\log|V_t| - \frac{\nu}{2}tr\left(V_t^{-1}RK_t\right), \tag{5.38}$$

where $d(k) = -\frac{k}{2}\log(2\pi)$ and $d(k,\nu) = \frac{\nu k}{2}\log(\nu) - \frac{\nu k}{2}\log(2) - \log\left(\Gamma_k\left(\frac{\nu}{2}\right)\right)$, and where Γ_k denotes the multivariate Gamma function such that $\Gamma_k\left(\frac{\nu}{2}\right) = \pi^{\frac{k(k-1)}{4}}\prod_{i=1}^{k}\Gamma\left(\frac{\nu}{2} + (1-i)/2\right)$.

To ensure positive definiteness of the covariance matrix, we need to consider some appropriate decomposition, for example Cholesky decomposition,

$$V_t = C_t C_t', \tag{5.39}$$

where C_t is lower triangular matrix with unique Cholesky factors. It follows from (5.39) that it suffices to specify a model for lower triangular part of C_t, thus our vector of parameters to be modelled is defined as

$$f_t = vech(C_t), \tag{5.40}$$

which results in a $k^* \times 1$ vector with $k^* = k(k+1)/2$. To operate the updating equation (5.32), we need the score vector and expected information matrix that we obtain as described in Section 5.3.3.

Proposition 4. *For the measurements densities (5.16) and (5.17), and for the decomposition (5.39), the score vector of dimension $k^* \times 1$ is given by*

$$\nabla_t = \frac{1}{2}\dot{V}_t' D_k' \big(V_t^{-1} \otimes V_t^{-1}\big) \left(\nu \big[vec(RK_t) - vec(V_t) \big] + \big[vec(\Lambda^{-1/2} r_t r_t' \Lambda^{-1/2}) - vec(V_t) \big] \right), \tag{5.41}$$

where $\dot{V}_t = L_k(I_{k^2} + K_k)(C_t \otimes I_k)\mathscr{L}_k'$.

Proof, see Appendix 5.A.1.

It follows from (5.24) and (5.25) that $\mathbb{E}[\nabla_t | \mathcal{F}_{t-1}] = 0$ when evaluated at true parameter values implying that ∇_t forms martingale sequence. The expression for the score in Proposition 4 indicates that when making one-step ahead prediction from V_t to V_{t+1}, information from the deviations of RK_t from the its mean V_t (which is $\mathcal{F}_{t-1}$ measurable) is ν times more relevant that information from deviations of $r_t r_t'$ from V_t (correcting for overnight variation Λ if r_t is vector of close-to-close returns). This model feature is pertinent as the outer-product of daily returns contains only a weak signal about the current covariance of assets as it does not exploit intraday information. It follows from Proposition 4 that GAS-based derivation of the model is not invariant to the decomposition applied to the covariance matrix V_t to ensure positive definiteness. The dependence enters through the term $\dot{V}_t$ which is unique for a selected decomposition. Note that $\dot{V}_t$ collects the first order derivatives of the full covariance matrix V_t with respect to the dynamic factors f_t.

Proposition 5. *For the measurements densities (5.16) and (5.17), and for the decomposition (5.39), the Fisher information matrix of dimension $k^* \times k^*$ is given by*

$$\mathcal{I}_t = \mathbb{E}[\nabla_t \nabla_t' | \mathcal{F}_{t-1}] = \frac{1+\nu}{4} \dot{V}_t' D_k' \big(V_t^{-1} \otimes V_t^{-1}\big) \big(I_{k^2} + K_k\big) D_k \dot{V}_t. \tag{5.42}$$

Proof, see Appendix 5.A.1.

The square root of the inverse of the conditional information matrix may be used to scale the score vector, such that $\mathbb{E}[s_t s_t'|\mathcal{F}_{t-1}] = I_{k^*}$ with s_t defined in (5.35). This scaling turns convenient as most of the well-known results and properties of the vector ARMA can be directly applied to (5.32) to obtain, for instance, conditions for stationarity or unconditional moments. However, scaling (5.35) implies the need to invert the Fisher information matrix $\mathcal{I}_t$ whose dimension grows at rate proportional to $O(k^2)$. This step is therefore the most computationally demanding.

The key distinguishing feature of our model specification is that each element of the innovation vector s_t exploits the full likelihood information. This feature turns relevant when avoiding the curse of dimensionality as we may consider diagonal specification defined through $B_i = diag(\beta_1^i, \ldots, \beta_{k^*}^i)$ and $A_j = diag(\alpha_1^j, \ldots, \alpha_{k^*}^j)$ or even a simple scalar version defined through $B_i = \beta^i I_{k^*}$ and $A_j = \alpha^j I_{k^*}$ in (5.32). In either case, the model dynamics allows for a complex interdependence between all variances and covariances such that the one-step update from V_t to V_{t+1} is driven by own as well as cross-asset effects. This feature is GAS specific and distinguishes our model from low frequency standard GARCH model driven by outer-product of daily returns $r_t r_t'$ and from the high frequency HEAVY model driven by realized measure, RK_t say, as both models require full parametrization to allow for cross-asset effects. Our model specification creates therefore some potential for reducing the number of static parameters to be estimated having still relatively rich relationships.

5.4 Simulation study

In this section, we provide simulation evidence on the statistical properties of the maximum likelihood estimation method for the model proposed in the previous section. We study estimation performance for varying sample size T and for varying dimension k. We start by discussing the estimation of static parameters in the model using the method of maximum likelihood. We then describe design of the Monte Carlo study and we discuss the simulation results.

5.4.1 Estimation

Given a set of T realizations of daily returns and covariance matrices of dimension k, the log-likelihood function is given by

$$\mathcal{L} = \sum_{t=1}^{T} \big(\mathcal{L}_{1,t} + \mathcal{L}_{2,t}\big), \tag{5.43}$$

where $\mathcal{L}_{1,t}$ and $\mathcal{L}_{2,t}$ are given in (5.37) and (5.38), respectively and where time-variation of V_t is determined by the recursion (5.32), decomposition (5.39) and parametrization (5.40).

The general GAS(p, q) recursion is parameterized with a finite-dimensional parameter vector defined by

$$\psi = \left[\omega', vec(B_1)', \ldots, vec(B_p)', vec(A_1)', \ldots, vec(A_q)', \nu, (\lambda_1, \ldots, \lambda_k)\right]', \tag{5.44}$$

which results in $(p+q)(k(k+1)/2)^2 + k(k+1)/2 + k + 1$ static parameters to be estimated and is of order $O(k^4)$. The recursion (5.32) needs to be initialized and it is natural to set $s_0 = 0$ and f_0 either to the unconditional first moment estimated from the data or it can be added to the vector of parameters ψ to be jointly estimated. For given parameter values ψ, the log-likelihood function can be evaluated in a straightforward way. In practice, ψ is unknown and estimation of all parameters is carried out by numerically maximizing (5.43) with respect to ψ. Maximization can be based on a standard quasi-Newton numerical optimization procedure with initial values of ψ determined from a basic grid search. The model estimations carried out in the simulation studies as well as in the empirical applications are based on the numerical derivatives. Derivation of analytical derivatives with respect to parameters collected in vector ψ, which would facilitate computation of robust standard errors, are left for future research.

With increasing dimension k, the estimation of model parameters may become computationally demanding. One approach to reduce the number of parameters to be estimated is based on covariance targeting as proposed by Engle and Mezrich (1996) for the GARCH models. Since the recursion (5.32) admits the vector ARMA structure, the model intercept can be expressed, if stationarity conditions are satisfied, using the unconditional moment,

$$f_{t+1} = (I_{k^*} - \sum_{i=1}^{p} B_i)\, vech(\mathbb{E}[f_t]) + \sum_{i=1}^{p} B_i f_{t-i+1} + \sum_{j=1}^{q} A_j s_{t-j+1}, \tag{5.45}$$

where $\mathbb{E}[f_t]$ is replaced by the sample moment estimator, $\widehat{\mathbb{E}[f_t]} = \widehat{C}$, with $T^{-1}\sum_{t=1}^{T} RK_t = \widehat{C}\widehat{C}'$ such that $\widehat{C}$ is the lower triangular matrix with Cholesky factors of a long run target as measured by the sample mean of the realized kernel series. Introduction of targeting makes model estimation be a two-step approach. We first reparametrize the model by substituting out the vector of constants using the unconditional mean which is replaced by a consistent estimator, we next maximize the log-likelihood. The targeting technique eliminates k^* parameters from the parameter space of ψ. To avoid the curse of dimensionality further reduction of the number of parameters can be achieved by setting $A_1, \ldots, A_q$ and $B_1, \ldots, B_p$ to diagonal matrices or to scalars. In either case, both diagonal and off-diagonal elements of the covariance matrix are driven by their own lagged values and, importantly, by the cross-asset effects. The cross-asset effects are captured by the individual elements of the score vector which all contain full likelihood information at each time update as discussed in Section 5.3.4.

5.4.2 Simulation results

We study properties of the likelihood-based estimation method by means of the simulation-based exercises. We simulate a series of $T \in \{250, 500, 1\,000\}$ daily returns and daily covariance matrices of dimension $k \in \{2, 5, 10\}$. For simplicity, we study the most parsimonious model specification with $p = q = 1$ in (5.32), and we consider the scalar formulation implied by $B = \beta I_{k^*}$, $A = \alpha I_{k^*}$ and $\Lambda = \lambda I_{k^*}$, with $k^* = k(k+1)/2$. The resulting Monte Carlo data generation process is here based on the parameter values,

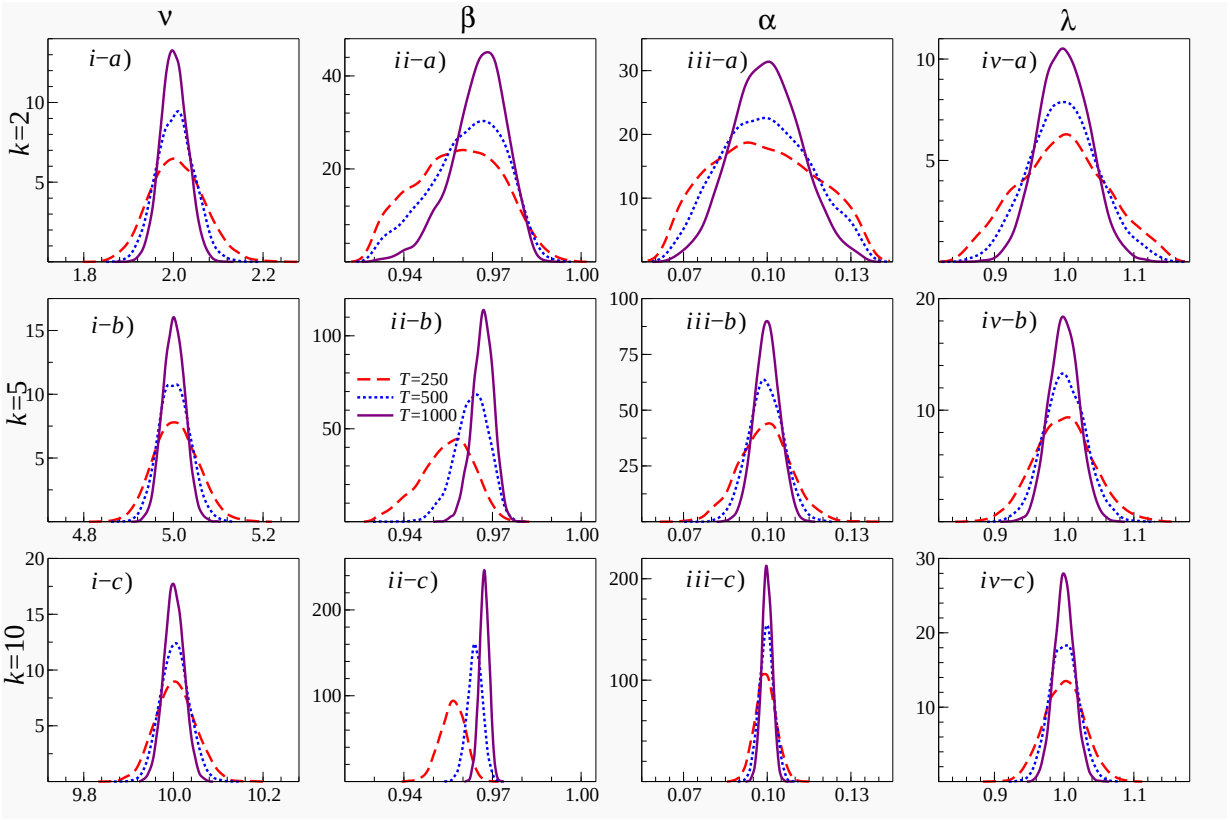
$$\nu = k, \qquad \omega = 0.10\ \iota_{k^*}, \qquad \beta = 0.97, \qquad \alpha = 0.10, \qquad \text{and} \qquad \lambda = 1, \tag{5.46}$$

where ι_{k^*} denotes column vector of ones of dimension k^*. The parameter values are in roughly line with the empirical estimates that we present in Section 5.5. The high value of the memory parameter β indicates that variances and covariances are persistent processes close to unit root what is found in the empirical studies. We simulate 5 000 datasets for our Monte Carlo study. For each generated dataset, we maximize the log-likelihood as discussed in the previous subsection, and we collect the likelihood-based estimates of parameters (5.46). We carry out unconstrained estimation and we consider covariance targeting.

The Monte Carlo results are graphically presented in Figure 5.1, where kernel estimates of the simulation densities of parameters are presented. Each panel of Figure 5.1 presents densities for the three values of T here considered. The obtained densities indicate that for increasing sample sizes T, the estimates concentrate more at their true values and densities are clearly more symmetric. There is some skewness and longer tails in the densities of the estimates for small sample size $T = 250$, particularly for the memory parameter β whose simulated density is skewed to the left. For this particular parameter, we report also some negative bias for small samples which is reflected by the shift of the mode of the simulation density to the left near $\beta = 0.97$. Precise estimation of this parameter requires relatively long time series. The number of degrees of freedom of the Wishart density ν appears to be estimated fairly robustly, even at moderate sample sizes, which might be surprising given that the model formulation is nonlinear in this parameter.

Given the multivariate model, the likelihood-based estimation depends not only the length of time series T but also on the dimension k. When we increase the dimension of assets in the simulation, we observe that with increasing k the shapes of densities improve as they are considerably more symmetric and peaked around their respective true values; compare panels *a*) to *c*). The improvement is particularly remarkable for parameters α and β. However, still to obtain a precise estimate of β the length of the time series is key requirement. To conclude, the maximum likelihood method works reasonably well in terms of estimating the model parameters and it provides evidence of the relative simplicity in empirically identifying the parameters that control the time-variation of the covariance matrices as we discuss in the next section.

Figure 5.1: *Simulation densities*



5.5 Empirical Illustrations

5.5.1 Dataset

Our data is constructed to capture variation across firms and across market conditions. The equities we analyze include some fifteen Dow Jones Industrial Average components with the following ticker symbols AA, AXP, BA, CAT, GE, HD, HON, IBM, JPM, KO, MCD, PFE, PG, WMT and XOM. The empirical study is based on data on consolidated trades (transaction prices) extracted from the Trade and Quote (TAQ) database through the Wharton Research Data Services (WRDS) system. The time-stamp precision is one second. The sample period spans ten years, from January 2, 2001 to December 31, 2010, with a total of $T = 2515$ trading days for all equities.

Before we construct a series of daily realized covariances and kernels as discussed in Section 5.2, we carry out some cleaning procedures to the raw transaction data. The importance of tick-by-tick data cleaning is highlighted by Brownless and Gallo (2006), Hansen and Lunde (2006) and Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009) who

provide a guideline on cleaning procedures based on the TAQ qualifiers that are included in the files and described in the TAQ User's Quide available at the WRDS website. In particular, we carry out the following steps: *i*) we delete entries with a time stamp outside the 9:30*am*-4*pm* window; *ii*) we delete entries with transaction price equal to zero; *iii*) we retain entries originating from a single exchange (NYSE in our application); *iv*) we delete entries with corrected trades (trades with a correction indicator, CORR $\neq 0$); *v*) we delete entries with abnormal sale condition (trades where COND has a letter code, except for "E" and "F"); *vi*) we use the median price for multiple transactions with the same time stamp; *vii*) we delete entries with prices that are above the ask plus the bid-ask spread.

We have in total 15 equities in our empirical study. To conserve space in our report, we provide results for some selected cases for $k \in \{2, 5, 15\}$. We randomly select ten 2-variate and ten 5-variate models. We present also results for the full 15-variate model.

In our empirical application we use the subsampled realized covariance and the realized kernel. The realized covariance is based on the calendar time sampling and we use returns sampled at $5min$ frequency which results in 79 price observations. By jittering the initial point we obtain 300 estimators each day that we average to get the subsampled measure; recall Section 5.2.1. The realized kernel makes use of refresh time sampling applied to k assets as discussed in Section 5.2.2.

Table 5.1 presents summary statistics on the number of observations and fraction of the data we retained by constructing refresh time sampling. Given the dimension k, we record the resulting daily number of price observations and daily p statistics defined in (5.8). These statistics are averaged over particular years in our sample. We observe that for the 2-variate system we retain on average of around 60-65% observations and this fraction is somewhat robust over time and across equities. The average number of refresh time observations is around 2800 and it moderatly varies in time with higher volatility during crisis in 2007-2009. Note that the average p statistics implies we have on average refresh observation every $8sec$ for $k = 2$.

For the 5×5 case the data loss is more pronounced. We retain around 35-40% and 1800 refresh observations on average. Again, the p statistics is rather flat in time. In this case, we have a vector of price observations on average around every $13sec$. For the full 15×15 case, the overall average of fraction of retained observations equals around 22%, while the average number of observations is around 950 implying refresh observation every $25sec$.

5.5.2 Capturing overnight variation

In our first empirical illustration, we wish to emphasize one empirical finding that concerns the difference between modeling open-to-close and close-to-close returns. While the realized measures are defined here over a period of a trading day i.e. from open-to-close, then the

Table 5.1: *Summary statistics for the refresh sampling scheme*

Equities	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
PANEL A										
2x2										
AA/CAT	803	1043	1340	1899	1919	2458	3538	3730	2810	2006
AXP/PFE	1805	2081	2486	2198	2372	2413	4007	4355	3527	2108
AXP/WMT	1508	1760	1865	2062	2323	2449	3994	4816	3900	2827
BA/HON	959	1248	1665	1719	2036	2171	3111	3069	2407	2154
CAT/KO	831	1144	1516	1934	2059	2382	3469	3809	3049	2585
GE/PFE	2064	2753	3061	3135	3156	3201	5105	5374	3514	1935
HD/JPM	1657	2022	2421	2329	2523	2817	4706	5454	3693	2906
IBM/PG	1566	1971	2390	2618	2659	3017	4252	4549	3493	2895
JPM/XOM	1476	1980	2516	2607	3044	3531	6187	7799	5747	4169
MCD/PG	1147	1516	1847	1969	2397	2517	3531	4330	3315	2442
5x5										
AA/AXP/IBM/JPM/WMT	827	940	1048	1304	1405	1553	2632	3074	2210	1526
AA/BA/CAT/GE/KO	570	736	933	1172	1247	1466	2340	2584	1790	1266
AXP/CAT/IBM/KO/XOM	671	885	1141	1272	1352	1520	2521	2787	2239	1924
BA/HD/JPM/PFE/PG	847	1060	1336	1332	1472	1639	2665	2920	2039	1395
BA/HD/MCD/PG/XOM	748	990	1232	1238	1462	1596	2483	2834	2009	1620
CAT/GE/KO/PFE/WMT	680	887	1055	1367	1481	1646	2625	2912	2070	1333
CAT/HON/IBM/MCD/WMT	626	783	951	1172	1332	1440	2186	2342	1857	1614
GE/IBM/JPM/PG/XOM	947	1256	1548	1586	1709	1915	3283	3773	2616	1863
HD/HON/KO/MCD/PG	662	868	1066	1136	1371	1414	2196	2443	1768	1432
HON/IBM/MCD/WMT/XOM	745	940	1079	1266	1537	1585	2408	2602	1994	1669
15x15										
AA/ ... /XOM	430	530	649	759	856	951	1613	1779	1267	894
PANEL B										
2x2										
AA/CAT	0.599	0.588	0.587	0.601	0.584	0.579	0.625	0.640	0.612	0.543
AXP/PFE	0.646	0.625	0.627	0.565	0.579	0.572	0.625	0.653	0.620	0.548
AXP/WMT	0.637	0.629	0.600	0.576	0.584	0.570	0.631	0.666	0.652	0.625
BA/HON	0.616	0.601	0.615	0.603	0.598	0.586	0.627	0.636	0.629	0.632
CAT/KO	0.583	0.573	0.577	0.595	0.584	0.585	0.636	0.651	0.643	0.626
GE/PFE	0.655	0.642	0.655	0.640	0.640	0.624	0.668	0.663	0.617	0.548
HD/JPM	0.644	0.625	0.635	0.615	0.621	0.607	0.652	0.636	0.575	0.582
IBM/PG	0.579	0.646	0.648	0.636	0.626	0.628	0.662	0.672	0.654	0.642
JPM/XOM	0.626	0.618	0.629	0.620	0.584	0.566	0.672	0.732	0.699	0.668
MCD/PG	0.643	0.628	0.624	0.610	0.621	0.597	0.634	0.662	0.643	0.637
5x5										
AA/AXP/IBM/JPM/WMT	0.338	0.338	0.322	0.347	0.354	0.348	0.396	0.407	0.371	0.329
AA/BA/CAT/GE/KO	0.314	0.288	0.308	0.336	0.334	0.339	0.385	0.394	0.375	0.334
AXP/CAT/IBM/KO/XOM	0.305	0.324	0.338	0.348	0.322	0.313	0.374	0.398	0.394	0.400
BA/HD/JPM/PFE/PG	0.357	0.348	0.363	0.345	0.354	0.354	0.400	0.395	0.360	0.328
BA/HD/MCD/PG/XOM	0.373	0.353	0.365	0.352	0.337	0.319	0.369	0.399	0.373	0.373
CAT/GE/KO/PFE/WMT	0.296	0.290	0.303	0.330	0.340	0.348	0.393	0.404	0.384	0.331
CAT/HON/IBM/MCD/WMT	0.305	0.326	0.330	0.333	0.339	0.336	0.382	0.396	0.385	0.388
GE/IBM/JPM/PG/XOM	0.358	0.366	0.384	0.371	0.361	0.352	0.416	0.426	0.392	0.362
HD/HON/KO/MCD/PG	0.359	0.347	0.354	0.353	0.362	0.348	0.393	0.405	0.385	0.389
HON/IBM/MCD/WMT/XOM	0.333	0.340	0.333	0.335	0.337	0.316	0.357	0.374	0.366	0.370
15x15										
AA/ ... /XOM	0.197	0.189	0.195	0.206	0.208	0.207	0.247	0.253	0.234	0.210

Note: Statistics are measured separately for each year in our sample. Panel A presents averages over the daily number of high-frequency observations maintained by the refresh sampling scheme. Panel B present averages of ratio of the data maintained by the refresh sampling scheme as measured by (5.8).

daily returns in the traditional GARCH modeling are defined as either open-to-close or close-to-close returns. In general, the interest is in the total variation including overnight effects, and a number of approaches have been proposed to capture it. Here, we provide

a model specification that should be sufficiently flexible to capture total variation where overnight effects are captured by the close-to-close returns. In case we use close-to-close returns, the individual elements of Λ in (5.23) should capture the additional variation due to the overnight effects. Using the univariate realized GARCH model Hansen et al. (2011) found that overnight variation may stand for around 25% of total daily variation, implying we could expect the individual elements of Λ equal around 1.25.

Table 5.2 presents maximum likelihood estimates for the selected portfolio of equities. We consider the model specification with $B = \beta I_{k^*}$ and $A = \alpha I_{k^*}$ and with covariance targeting. One remarkable observation is that the parameter estimates of ν, α and β do not display large variation across different equity pairs. However, we observe that estimates differ in the values of λs dependently on the definition of returns. For the open-to-close returns the values of individual elements of Λ are close to unity and somewhat below. Instead, if we model the close-to-close returns the individual elements of Λ are somewhere in between 1.20-1.40. These results suggest that if the interest is also to predict daily total variation of assets, then a simple link (5.23) may turn convenient and useful for capturing overnight variation. These results also indicate that a more parsimonious model may be obtained by restricting the coefficients of Λ to $\Lambda = \lambda I_{k^*}$ in a similar fashion as we restricted $A = \alpha I_{k^*}$ and $B = \beta I_{k^*}$ (the more general results indicate that these parameters are virtually the same for different equities). It is also seen that estimates of other model parameters ν, α and β are virtually the same. In general, we find estimates of β being close to unity implying high persistence of the covariance coefficients. We observe also that dynamics of V_t put more weight on realized kernel measures which is expressed in the estimates of ν which magnify the information coming from realized measures relative to outer-product of daily returns; recall Proposition 4.

5.5.3 Benchmarking against common alternatives

We now compare performance of our model against some widely applied alternatives. We relate our model to the so-called BEKK model of Engle and Kroner (1995). In the standard BEKK$(1,1)$ model it is assumed $r_t|\mathcal{F}_{t-1} \sim \mathcal{N}(0, V_t)$. The covariance of assets is assumed to be driven by outer-product of daily returns,

$$V_{t+1} = CC' + BV_tB' + Ar_tr_t'A', \tag{5.47}$$

where A, B, and C are $k \times k$ matrices with static parameters of the model, and C is lower triangular. The scalar BEKK with covariance targeting takes the form,

$$V_{t+1} = (1 - b - a)\widehat{R} + bV_t + ar_tr_t', \qquad a, b \geq 0, \tag{5.48}$$

where $\widehat{R} = T^{-1}\sum_{t=1}^{T} r_tr_t'$ is the sample covariance of daily returns. When contrasting the BEKK model to the result in Proposition 4, it is evident that the BEKK specification

Table 5.2: *Maximum likelihood estimates for open-to-close and close-to-close returns*

	open-to-close									close-to-close								
ν	β	α	λ_1	λ_2	λ_3	λ_4	λ_5	*log-lik*		ν	β	α	λ_1	λ_2	λ_3	λ_4	λ_5	*log-lik*
2x2																		
11.6133	0.9787	0.0845	0.9026	0.9697	-	-	-	-20543.62		11.6436	0.9785	0.0786	1.3505	1.4624	-	-	-	-21556.16
(0.1760)	(0.0019)	(0.0026)	(0.0249)	(0.0267)	-	-	-	-		(0.1763)	(0.0019)	(0.0025)	(0.0373)	(0.0401)	-	-	-	-
9.7035	0.9788	0.0788	1.0324	0.8887	-	-	-	-17697.90		9.7497	0.9784	0.0731	1.5314	1.2973	-	-	-	-18649.30
(0.1449)	(0.0017)	(0.0028)	(0.0293)	(0.0252)	-	-	-	-		(0.1458)	(0.0017)	(0.0027)	(0.0438)	(0.0368)	-	-	-	-
10.6768	0.9822	0.0745	1.0454	0.8049	-	-	-	-15913.94		10.7320	0.9829	0.0684	1.5445	1.1411	-	-	-	-16821.31
(0.1606)	(0.0014)	(0.0025)	(0.0293)	(0.0225)	-	-	-	-		(0.1615)	(0.0013)	(0.0023)	(0.0435)	(0.0319)	-	-	-	-
10.0744	0.9627	0.0968	0.9407	0.8436	-	-	-	-18192.85		10.0948	0.9627	0.0956	1.3274	1.1562	-	-	-	-19013.12
(0.1507)	(0.0026)	(0.0029)	(0.0257)	(0.0230)	-	-	-	-		(0.1509)	(0.0026)	(0.0028)	(0.0358)	(0.0312)	-	-	-	-
11.9814	0.9722	0.0679	1.0118	0.8491	-	-	-	-14579.31		12.0297	0.9695	0.0683	1.5234	1.1390	-	-	-	-15446.25
(0.1816)	(0.0022)	(0.0020)	(0.0287)	(0.0240)	-	-	-	-		(0.1820)	(0.0023)	(0.0021)	(0.0431)	(0.0321)	-	-	-	-
9.9600	0.9679	0.0867	0.8936	0.8609	-	-	-	-16177.99		9.9907	0.9662	0.0872	1.3179	1.2485	-	-	-	-17120.73
(0.1487)	(0.0023)	(0.0029)	(0.0249)	(0.0241)	-	-	-	-		(0.1489)	(0.0024)	(0.0031)	(0.0365)	(0.0349)	-	-	-	-
11.2892	0.9734	0.1117	0.8913	0.9554	-	-	-	-19021.08		11.3410	0.9732	0.1137	1.2511	1.2624	-	-	-	-19777.63
(0.1702)	(0.0019)	(0.0032)	(0.0246)	(0.0263)	-	-	-	-		(0.1704)	(0.0019)	(0.0032)	(0.0343)	(0.0347)	-	-	-	-
10.5581	0.9538	0.0657	0.9803	0.7759	-	-	-	-11769.84		10.5793	0.9529	0.0675	1.3972	1.0822	-	-	-	-12627.73
(0.1587)	(0.0026)	(0.0021)	(0.0276)	(0.0218)	-	-	-	-		(0.1587)	(0.0026)	(0.0021)	(0.0394)	(0.0305)	-	-	-	-
11.5582	0.9740	0.0977	0.9879	0.8906	-	-	-	-16722.57		11.5929	0.9740	0.0995	1.2947	1.1748	-	-	-	-17398.08
(0.1744)	(0.0017)	(0.0026)	(0.0274)	(0.0248)	-	-	-	-		(0.1745)	(0.0017)	(0.0026)	(0.0358)	(0.0324)	-	-	-	-
9.5205	0.9586	0.0609	0.8401	0.7709	-	-	-	-13099.50		9.5306	0.9602	0.0596	1.1440	1.0641	-	-	-	-13887.90
(0.1420)	(0.0026)	(0.0020)	(0.0240)	(0.0220)	-	-	-	-		(0.1420)	(0.0025)	(0.0020)	(0.0327)	(0.0302)	-	-	-	-
5x5																		
16.8415	0.9827	0.0535	0.9239	0.9636	0.9603	0.9155	0.7943	-46208.96		16.8776	0.9825	0.0530	1.3585	1.4138	1.3638	1.2146	1.1441	-48380.22
(0.1081)	(0.0007)	(0.0009)	(0.0249)	(0.0244)	(0.0249)	(0.0231)	(0.0210)	-		(0.1083)	(0.0007)	(0.0009)	(0.0363)	(0.0355)	(0.0354)	(0.0306)	(0.0304)	-
16.8578	0.9805	0.0511	0.9058	0.9627	0.9721	0.8749	0.8650	-46256.39		16.8775	0.9801	0.0507	1.3133	1.3545	1.4445	1.2521	1.1583	-48440.87
(0.1082)	(0.0008)	(0.0008)	(0.0240)	(0.0257)	(0.0253)	(0.0225)	(0.0236)	-		(0.1083)	(0.0008)	(0.0008)	(0.0346)	(0.0358)	(0.0373)	(0.0319)	(0.0314)	-
17.6635	0.9788	0.0477	1.0048	0.9809	0.9935	0.8541	0.8940	-35028.57		17.7017	0.9782	0.0476	1.4498	1.4786	1.4080	1.1609	1.1899	-37140.69
(0.1142)	(0.0008)	(0.0007)	(0.0263)	(0.0258)	(0.0259)	(0.0231)	(0.0235)	-		(0.1143)	(0.0008)	(0.0007)	(0.0378)	(0.0387)	(0.0367)	(0.0313)	(0.0311)	-
16.4246	0.9753	0.0532	0.9648	0.9044	0.9482	0.8822	0.7941	-44412.52		16.4424	0.9751	0.0539	1.3640	1.2573	1.2552	1.2762	1.0787	-46438.44
(0.1051)	(0.0009)	(0.0009)	(0.0261)	(0.0241)	(0.0250)	(0.0238)	(0.0214)	-		(0.1051)	(0.0009)	(0.0009)	(0.0367)	(0.0333)	(0.0331)	(0.0344)	(0.0290)	-
17.0484	0.9698	0.0474	0.9713	0.9332	0.8654	0.7912	0.9069	-38077.34		17.0658	0.9702	0.0473	1.3681	1.2917	1.1760	1.0813	1.2035	-40026.13
(0.1097)	(0.0010)	(0.0008)	(0.0263)	(0.0250)	(0.0235)	(0.0213)	(0.0242)	-		(0.1097)	(0.0010)	(0.0007)	(0.0368)	(0.0345)	(0.0319)	(0.0290)	(0.0319)	-
17.0002	0.9762	0.0469	1.0033	0.8877	0.8577	0.8887	0.8185	-35613.46		17.0311	0.9753	0.0475	1.4928	1.2941	1.1723	1.2857	1.1671	-37847.55
(0.1093)	(0.0009)	(0.0008)	(0.0268)	(0.0227)	(0.0231)	(0.0239)	(0.0217)	-		(0.1093)	(0.0010)	(0.0009)	(0.0396)	(0.0329)	(0.0316)	(0.0346)	(0.0310)	-
16.1954	0.9721	0.0477	0.9935	0.8556	0.9855	0.8930	0.8115	-39998.86		16.2229	0.9715	0.0479	1.4809	1.1891	1.3951	1.2165	1.1667	-42156.51
(0.1034)	(0.0010)	(0.0008)	(0.0261)	(0.0222)	(0.0258)	(0.0244)	(0.0216)	-		(0.1035)	(0.0010)	(0.0008)	(0.0387)	(0.0305)	(0.0365)	(0.0331)	(0.0310)	-
17.4267	0.9729	0.0556	0.8480	0.9757	0.9223	0.7754	0.8804	-33329.28		17.4430	0.9727	0.0564	1.2299	1.3754	1.2166	1.0635	1.1664	-35307.46
(0.1123)	(0.0008)	(0.0008)	(0.0214)	(0.0252)	(0.0239)	(0.0207)	(0.0231)	-		(0.1123)	(0.0008)	(0.0008)	(0.0308)	(0.0355)	(0.0315)	(0.0283)	(0.0304)	-
15.9711	0.9682	0.0479	0.9218	0.8841	0.8465	0.8720	0.7953	-36269.75		15.9865	0.9686	0.0478	1.2772	1.2318	1.1476	1.1917	1.0797	-38233.17
(0.1017)	(0.0011)	(0.0008)	(0.0246)	(0.0236)	(0.0227)	(0.0237)	(0.0212)	-		(0.1018)	(0.0010)	(0.0008)	(0.0340)	(0.0327)	(0.0306)	(0.0323)	(0.0287)	-
16.9796	0.9716	0.0471	0.8737	0.9976	0.8909	0.8162	0.9188	-35295.14		17.0007	0.9713	0.0475	1.2125	1.3977	1.2087	1.1720	1.2186	-37297.00
(0.1092)	(0.0010)	(0.0007)	(0.0229)	(0.0261)	(0.0244)	(0.0217)	(0.0244)	-		(0.1092)	(0.0010)	(0.0007)	(0.0316)	(0.0365)	(0.0329)	(0.0312)	(0.0322)	-

ν	β	α	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}	*log-lik*
15x15																		
open-to-close																		
28.3960	0.9828	0.0268	0.9196	0.9741	0.9755	0.9716	0.8328	0.8907	0.8603	0.9900	0.9101	0.8642	0.8939	0.8962	0.8031	0.8120	0.9032	-69226.50
(0.0572)	(0.0003)	(0.0002)	(0.0235)	(0.0237)	(0.0251)	(0.0241)	(0.0188)	(0.0221)	(0.0200)	(0.0246)	(0.0220)	(0.0225)	(0.0237)	(0.0235)	(0.0209)	(0.0203)	(0.0227)	-
close-to-close																		
28.4060	0.9827	0.0269	1.3078	1.3415	1.3511	1.4280	1.1993	1.2406	1.1426	1.4063	1.2030	1.1720	1.1958	1.2751	1.0857	1.1624	1.1677	-75245.35
(0.0572)	(0.0003)	(0.0002)	(0.0331)	(0.0322)	(0.0342)	(0.0351)	(0.0270)	(0.0307)	(0.0262)	(0.0349)	(0.0291)	(0.0304)	(0.0316)	(0.0332)	(0.0280)	(0.0292)	(0.0290)	-

Note: The particular combination of equities for the case $k \in \{2, 5\}$ are in the same order as in the Table 5.1. Standard errors are shown in parentheses.

does not exploit high-frequency data to infer about the current level of covariances. Also, the restricted scalar specification of the BEKK formulation does not capture cross-asset effects.

We consider also an Exponentially Weighted Moving Average (EWMA) approach applied to the high-frequency based realized kernels. The EMWA filter is often used by practitioners and regulators, e.g. in the RiskMetrics of J.P.Morgan (1996). The EMWA assumes that the covariance matrix is an integrated process given by

$$V_{t+1} = bV_t + (1-b)RK_t, \tag{5.49}$$

governed by a fixed smoothing parameter b which is typically set equal to 0.96.

The models here studied are non-nested, thus direct log-likelihood ratio tests are not possible. We evaluate the performance of the GAS model relative to the BEKK and EMWA using some two loss functions. We use a quasi-likelihood loss function given by

$$Q(V_t, \Sigma_t) = \log |V_t| + tr\left(V_t^{-1}\Sigma_t\right), \tag{5.50}$$

and we use the root mean squared error based on the matrix norm given by

$$F(V_t, \Sigma_t) = ||\Sigma_t - V_t||^{1/2} = \left[\sum_{i,j}(\Sigma_{ij,t} - V_{ij,t})^2\right]^{1/2}, \tag{5.51}$$

where Σ_t is true latent covariance matrix and V_t is model-based covariance matrix. In our application, the proxy for latent covariance matrix is based on the realized kernel and on the subsampled realized covariance using $5min$ returns. As the in-sample fit, we report the average loss functions. In this exercise, we study the open-to-close returns and we use baseline $p = q = 1$ model specification.

Table 5.3 presents loss functions for the selected cases when $k = 2$ and $k = 5$, and for the case when variation of all equities is jointly modeled, i.e. when $k = 15$. The obtained results indicate that the our developed model has overall substantially better fit than the alternative BEKK and EWMA models. The difference in the performance relative to the BEKK model can be substantial with improvement in in-sample fit as much as 10% or even 15%. This improvement is a direct consequence of incorporating the high-frequency based realized measures to build the model for daily covariance. Since the EWMA model is based here on realized measures, it corresponds closer to the performance of our model.

Figure 5.2 presents an example of model-based covariances for the AA/CAT pair. When relating the model-based estimates to the subsampled realized covariance, we observe a close match of the GAS-based model. The GAS-based covariances appear to adapt rather quickly to changes in the level of covariances. In contrast, the traditional BEKK model smooths the outer-product of daily returns without exploiting high-frequency information, and the response to changes is remarkably longer as seen in panels *i-b*) and *ii-b*) which zoom

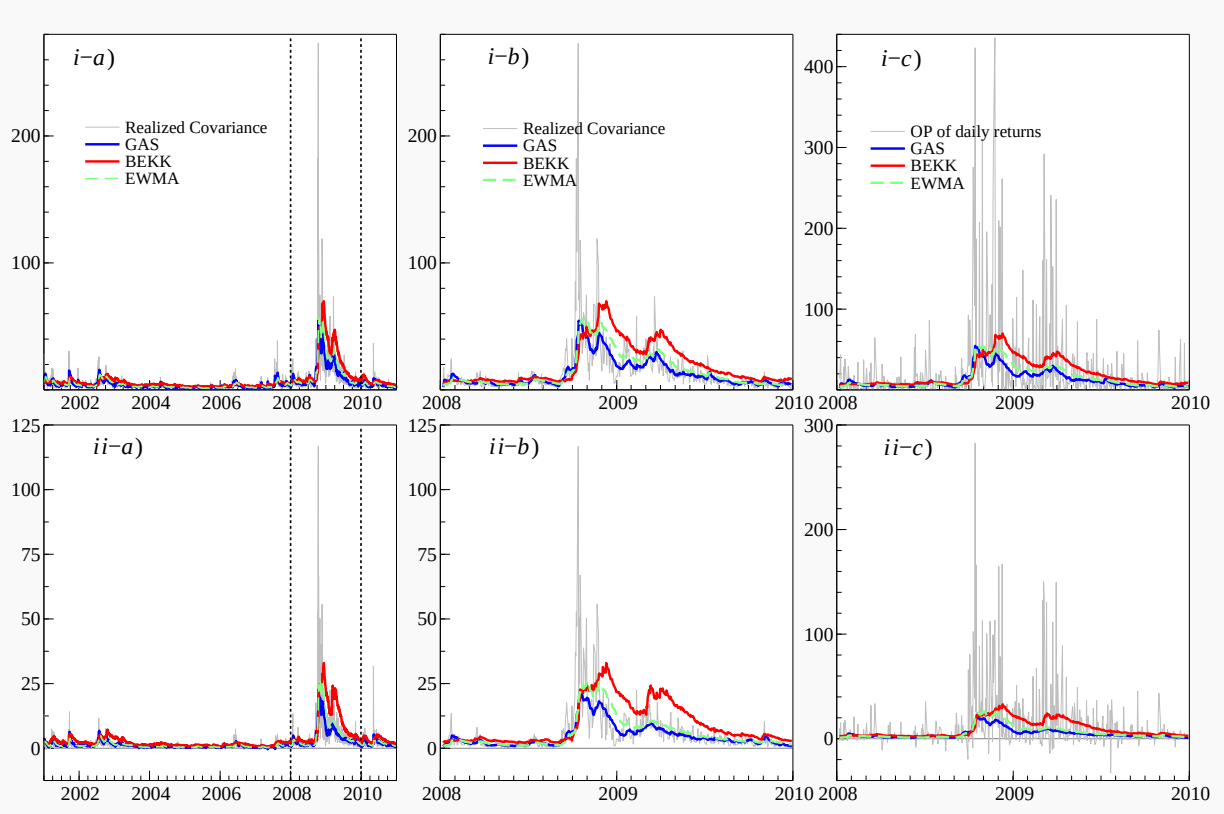
Table 5.3: *Values of loss functions*

Model	GAS				BEKK				EWMA			
Proxy	RK		RC		RK		RC					
Equities	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F
2x2												
AA/CAT	4.1733	3.4624	4.1492	3.3865	4.3524	4.1861	4.3310	4.1454	4.2374	3.9384	4.2085	3.8783
AXP/PFE	3.4868	3.1494	3.4530	3.0417	3.6947	3.8117	3.6431	3.7033	3.5523	3.5001	3.5107	3.4015
AXP/WMT	3.2134	2.9889	3.1728	2.8841	3.4367	3.6223	3.3795	3.5111	3.2853	3.3840	3.2399	3.2936
BA/HON	3.6423	2.7597	3.6312	2.6717	3.8234	3.0998	3.8094	3.0182	3.7297	3.0990	3.7066	2.9841
CAT/KO	3.0124	1.8930	2.9797	1.8130	3.1466	2.1675	3.1047	2.0885	3.0848	2.1736	3.0463	2.1013
GE/PFE	3.2411	2.6894	3.2050	2.6442	3.4208	2.8632	3.3629	2.8047	3.3119	2.9821	3.2674	2.9383
HD/JPM	3.8638	3.9762	3.8241	3.9342	4.1177	5.0781	4.0541	5.0280	3.9630	4.5966	3.9187	4.5604
IBM/PG	2.4108	1.7714	2.3595	1.7350	2.6516	1.6967	2.6016	1.6342	2.4889	1.7055	2.4398	1.6707
JPM/XOM	3.4104	3.5625	3.3827	3.5474	3.6305	4.4573	3.6073	4.4568	3.5315	4.1881	3.5027	4.1810
MCD/PG	2.6451	1.7553	2.6236	1.7095	2.8919	1.7174	2.8902	1.6696	2.7084	1.7619	2.6893	1.7106
5x5												
AA/BA/CAT/GE/KO	8.1170	6.1703	8.2090	5.9361	8.6177	7.2624	8.7292	7.1123	8.2465	6.7033	8.3239	6.4961
AXP/CAT/IBM/KO/XOM	6.9079	5.2938	6.9762	5.0869	7.5404	6.4618	7.6134	6.3033	7.0751	5.7260	7.1339	5.5476
BA/HD/JPM/PFE/PG	7.8779	6.1979	7.9521	5.9913	8.5163	7.2225	8.6190	7.0711	8.0156	6.4684	8.0846	6.2925
BA/HD/MCD/PG/XOM	7.2261	4.6157	7.2995	4.4767	7.7753	4.9079	7.8867	4.8074	7.3740	4.7933	7.4471	4.6741
CAT/GE/KO/PFE/WMT	6.9424	4.7605	7.0368	4.5630	7.5198	5.3290	7.6032	5.1539	7.0663	5.1475	7.1447	4.9590
CAT/HON/IBM/MCD/WMT	7.3463	5.0141	7.4483	4.7648	7.9266	5.6075	8.0379	5.4017	7.4813	5.3059	7.5669	5.0601
GE/IBM/JPM/PG/XOM	6.7398	6.5137	6.7637	6.3736	7.5618	7.4956	7.5772	7.4019	6.9224	6.6809	6.9389	6.5781
HD/HON/KO/MCD/PG	6.9314	4.7418	7.0472	4.5186	7.4972	5.0000	7.6587	4.8156	7.0671	4.8757	7.1754	4.6687
HON/IBM/MCD/WMT/XOM	6.9010	4.5943	6.9908	4.4136	7.4820	5.0771	7.5793	4.9153	7.0623	4.9146	7.1387	4.7307
15x15												
AA/.../XOM	20.5015	16.2492	21.5056	15.5803	23.2455	19.3341	24.4672	18.9041	20.7784	16.4803	21.7681	15.8529

Note: In proxy input, RK stands for realized kernel while RC for the subsampled realized covariance using $5min$ returns.

on the high volatile period in 2008-2009 years. It is worth noting that since our model formulation weights more the realized measures that outer-products of daily returns, it is seen that GAS-based covariance signal remains somewhat robust to large return innovations; see panels *i-c*) and *ii-c*).

Figure 5.2: *Model-based (co)variances*



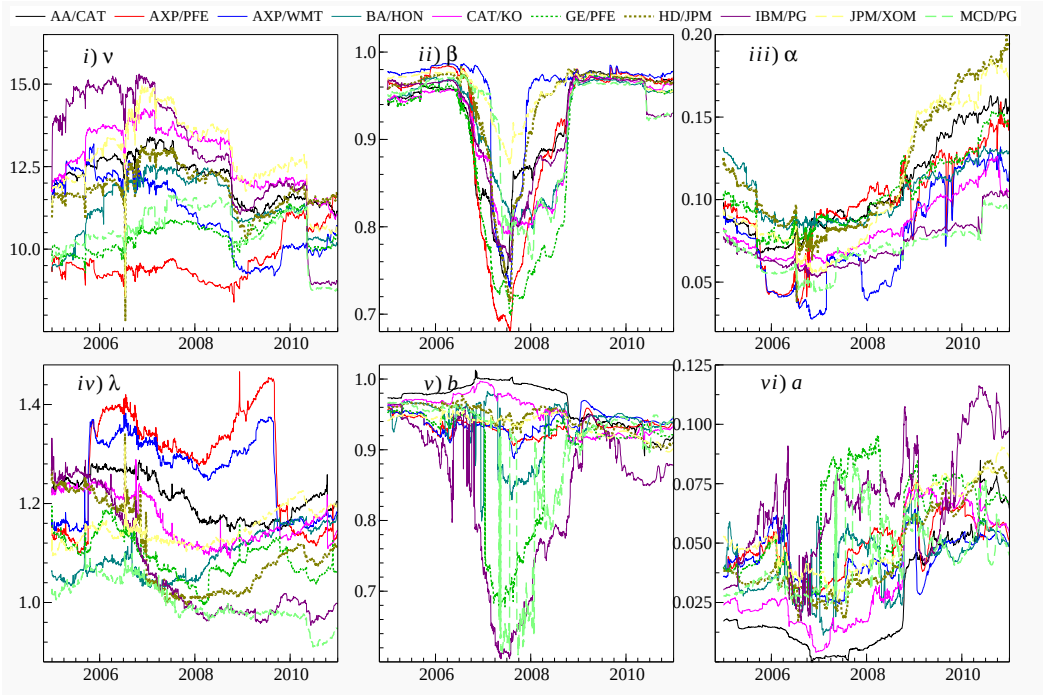
Note: Panels *i-a*) to *i*)-*c*) plot the case of variance of AA equity, while panels *ii-a*) to *ii*)-*c*) present the case of covariance of pair AA/CAT.

5.5.4 Parameter stability

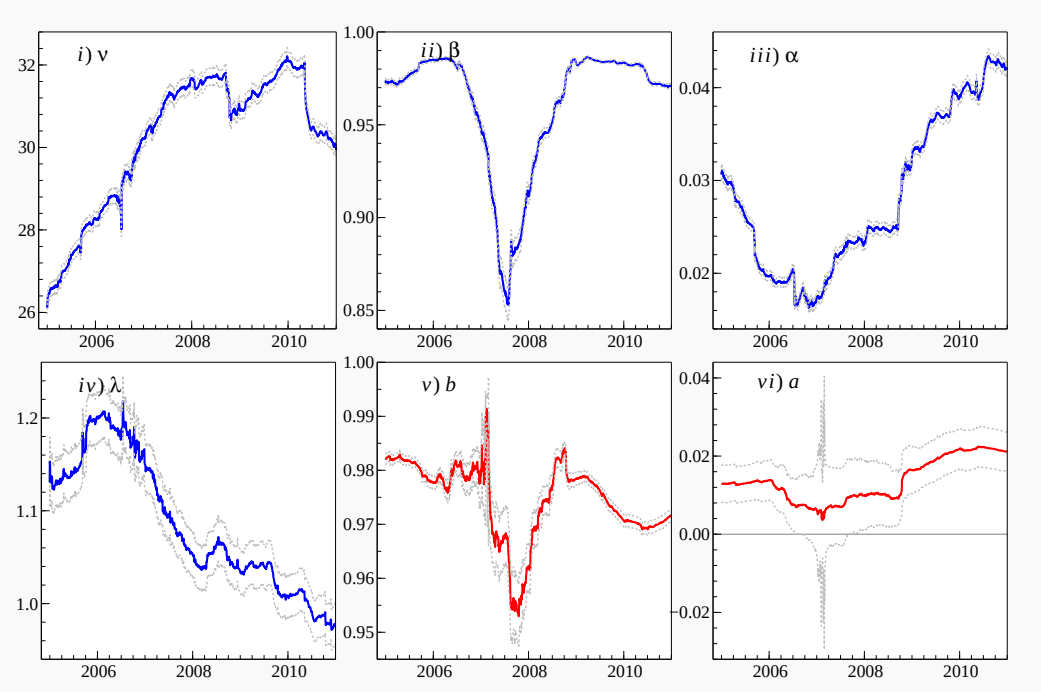
We now turn our attention to the stability of the model parameters. To address this concern, we focus on the most parsimonious scalar specification of the model. We use the close-to-close returns and restrict $\Lambda = \lambda I_{k^*}$ driven by the findings from Section 5.5.2. We estimate the model parameters of both GAS and BEKK models using the quasi-likelihood method based on a moving window of a period of four years of data implied by our choice $T = 1000$. We record the maximum likelihood estimates at the time of the last data entry in the moving sample. We carry out this exercise for a dimension of $k = 2$ and $k = 15$.

Figure 5.3a presents moving window parameter estimates for the bivariate models with

Figure 5.3: Moving window parameter estimates



(a) The figure displays model parameters obtained from moving window estimation of length $T = 1000$ and $k = 2$.



(b) The figure displays model parameters obtained from moving window estimation of length $T = 1000$ and $k = 15$.

the choice of equities being the same as in Table 5.1; Figure 5.3b presents moving window parameter estimates for the full 15×15 case. When looking at the moving window estimates of the persistence parameter β and b we observe they behave somewhat similar, decreasing during 2007-2008. However, the variation in the moving window estimates of b seems to be more dramatic, likewise in the moving window estimates of a. The more apparent instability of parameters in the BEKK model might result from occurrences of large price changes during the crisis. Note that the our model only partially exploits the information from daily returns when modeling covariance matrices. The more stable moving window estimates of model parameters of the GAS model relative to the BEKK model are also seen when estimating the full model; see Figure 5.3b where we also display associated 95% confidence intervals based on the estimates of standard errors of parameters. The more apparent instability of the BEKK parameters is there reflected in the more erratic confidence intervals.

5.6 Summary and Conclusions

In this chapter we propose a new formulation for the purpose of modeling the covariance matrix of financial assets. The modeling strategy is chosen to capture complex temporal interdependencies among variances and covariances of assets. There are two distinguishing features of our model as compared to the alternative frameworks. First, when making the one-step ahead prediction of the covariance matrix of assets, we make use of both low and high-frequency data. The information from the two frequencies is contained in the signal about the current level of volatilities and covariances. The high-frequency based measure receives most weight given that it exploits intraday data of financial assets to infer about covariances. Second, we show that each innovation driving the update of covariance matrix exploits full likelihood information. Consequently, even a simple scalar specification allows for complex interdependences between variances and covariances of all assets under consideration. Our model can therefore attain parsimonious formulation which is a convenient property for multivariate volatility models. In an application to a collection of assets from the NYSE exchange, we compare our model to the GARCH and EWMA models. We show that in-sample fit of our model generally dominates the fit from alternatives. This result is consistent across different subsets of equities but also when using the whole set of assets. The proposed model is also capable to track sudden changes in volatility and the dependence structure of the assets in a more efficient way than the standard multivariate GARCH model.

5.A Appendix

5.A.1 Proofs

We use the matrix notations and definitions as described in Section 5.3.1. To proof our results in this article, we make use of the following matrix calculus results. For a $p \times q$ matrix function $F(X)$ and a $m \times n$ matrix of variables X, the derivative of $F(X)$ with respect to $x = vec(X)$, denoted by the $pq \times mn$ matrix $DF(X)$, is given by

$$DF(X) = \frac{\partial\, vec(F(X))}{\partial x'}. \tag{5A.1}$$

If $F(X)$ is $m \times m$ symmetric matrix, we have

$$\frac{\partial\, vec(F(X))}{\partial x'} = D_m \frac{\partial\, vech(F(X))}{\partial x'}. \tag{5A.2}$$

The intermediate results for any $k \times k$ matrix are

$$\begin{aligned} \frac{\partial \log |AXB|}{\partial X} &= (vec((X^{-1})'))', \qquad (A, B \quad nonsingular), \\ \frac{\partial X^{-1}}{\partial X} &= -((X^{-1})' \otimes X^{-1}), \\ \frac{\partial\, tr(AXB)}{\partial X} &= A'B'. \end{aligned} \tag{5A.3}$$

The above results are combined with the following result for any $k \times k$ matrices A, B and C with B being symmetric matrix

$$vec(ABC) = (C' \otimes A)\, vec(B). \tag{5A.4}$$

Proof of Proposition 4: We derive the score vector whose general form is given by (5.33). It follows from (5.37) and (5.38) that the relevant parts of log-likelihoods for the score vector derivation are given by

$$\mathcal{L}_{1,t} = -\frac{1}{2}\Big(\log |\Lambda^{1/2} V_t \Lambda^{1/2}| + tr\big((\Lambda^{1/2} V_t \Lambda^{1/2})^{-1} r_t r_t'\big)\Big), \tag{5A.5}$$

$$\mathcal{L}_{2,t} = -\frac{\nu}{2}\Big(\log |V_t| + tr\big(V_t^{-1} R K_t\big)\Big). \tag{5A.6}$$

We consider the Cholesky decomposition (5.39) of the covariance matrix V_t and our GAS vector of parameters f_t is given by (5.40). Using the chain rule for vector differentiation, we can write the score for individual measurement densities (5.16) and (5.17) as given by

$$\frac{\partial \log\, \varphi_i(Z_t^i | f_t, \mathcal{F}_{t-1}; \psi)}{\partial f_t'} = \frac{\partial \log\, \varphi_i(Z_t^i | f_t, \mathcal{F}_{t-1}; \psi)}{\partial (vec(V_t))'} \frac{\partial\, vec(V_t)}{\partial\, vech(V_t)'} \frac{\partial\, vech(V_t)}{\partial f_t'}. \tag{5A.7}$$

We first differentiate the measurement density of returns (5A.5). Using (5A.3) and (5A.4), and the fact that V_t is symmetric matrix, we obtain

$$\begin{aligned}
\frac{\partial \mathcal{L}_{1,t}}{\partial\, vec(V_t)'} &= -\frac{1}{2}\Big((vec(V_t^{-1}))' - (vec(\Lambda^{-1/2} r_t r_t' \Lambda^{-1/2}))'(V_t^{-1} \otimes V_t^{-1})\Big) \\
&= -\frac{1}{2}\Big((vec(V_t^{-1} V_t V_t^{-1}))' - (vec(\Lambda^{-1/2} r_t r_t' \Lambda^{-1/2}))'(V_t^{-1} \otimes V_t^{-1})\Big) \\
&= -\frac{1}{2}\Big((vec(V_t))'(V_t^{-1} \otimes V_t^{-1}) - (vec(\Lambda^{-1/2} r_t r_t' \Lambda^{-1/2}))'(V_t^{-1} \otimes V_t^{-1})\Big) \\
&= \frac{1}{2}\Big((vec(\Lambda^{-1/2} r_t r_t' \Lambda^{-1/2}))' - (vec(V_t))'\Big)(V_t^{-1} \otimes V_t^{-1}), \qquad (5A.8)
\end{aligned}$$

and similarly for the measurement density of covariance (5A.6), we have

$$\begin{aligned}
\frac{\partial \mathcal{L}_{2,t}}{\partial\, vec(V_t)'} &= -\frac{\nu}{2}\Big((vec(V_t^{-1}))' - (vec(RK_t))'(V_t^{-1} \otimes V_t^{-1})\Big) \\
&= -\frac{\nu}{2}\Big((vec(V_t^{-1} V_t V_t^{-1}))' - (vec(RK_t))'(V_t^{-1} \otimes V_t^{-1})\Big) \\
&= -\frac{\nu}{2}\Big((vec(V_t))'(V_t^{-1} \otimes V_t^{-1}) - (vec(RK_t))'(V_t^{-1} \otimes V_t^{-1})\Big) \\
&= \frac{\nu}{2}\Big((vec(RK_t))' - (vec(V_t))'\Big)(V_t^{-1} \otimes V_t^{-1}). \qquad (5A.9)
\end{aligned}$$

The result for $\dot{V}_t = \frac{\partial\, vech(V_t)}{\partial f_t'}$ is based on the following Theorem.

Theorem 5.A.1. *Suppose V is a symmetric positive definite $k \times k$ matrix and C is a lower triangular $k \times k$ matrix such that $V = CC'$. Then $vech(V)$ is a totally differentiable function of $vech(C)$ and*

$$\dot{V} = \frac{\partial\, vech(V)}{\partial\, vech(C)} = L_k(I_{k^2} + K_k)(C \otimes I_k)\, \mathscr{L}_k'\,. \qquad (5A.10)$$

Proof of Theorem 5.A.1: The proof is based on the rules for matrix manipulations stated in Magnus and Neudecker (1988). Since $V = CC'$, we obtain

$$\mathrm{d}V = (\mathrm{d}C)C' + C(\mathrm{d}C'). \qquad (5A.11)$$

Vectorizing and using the result that $vec(\mathrm{d}X)A = (A' \otimes I)\mathrm{d}vecX$ (Magnus and Neudecker (1988, p.182)), we have

$$\begin{aligned}
\mathrm{d}\, vec(V) &= (C \otimes I_k)\mathrm{d}\, vec(C) + (I_k \otimes C)\mathrm{d}\, vec(C'), \\
&= (C \otimes I_k)\mathrm{d}\, vec(C) + (I_k \otimes C)K_k \mathrm{d}\, vec(C), \\
&= \big((C \otimes I_k) + (I_k \otimes C)K_k\big)\mathrm{d}\, vec(C). \qquad (5A.12)
\end{aligned}$$

Since for any $k \times k$ matrices A and B, we have $K_k(A \otimes B) = (B \otimes A)K_k$ (see Magnus and Neudecker (1988, p.46) for more general rules on commutation matrix), we have

$$\begin{aligned} \mathrm{d}\, vec(V) &= \big((C \otimes I_k) + K_k(C \otimes I_k)\big)\mathrm{d}\, vec(C), \\ &= \big((I_{k^2} + K_k)(C \otimes I_k)\big)\mathrm{d}\, vec(C). \end{aligned} \tag{5A.13}$$

Finally, left-multiplying by L_k gives

$$\mathrm{d}\, vech(V) = L_k\big((I_{k^2} + K_k)(C \otimes I_k)\big)\, \mathscr{L}'_k\, \mathrm{d}\, vech(C), \tag{5A.14}$$

where

$$\mathrm{d}\, vec(C) = \mathscr{D}_k\, \mathrm{d}\, vech(C) = \mathscr{L}'_k\, \mathrm{d}\, vech(C) \tag{5A.15}$$

because C is lower triangular matrix. □

The results in (5A.8), (5A.9) and the result of Theorem 5.A.1 combined with (5.33) complete the proof of Proposition 4. □

Proof of Proposition 5: We derive the Fisher information matrix whose general form is given by (5.34). Using the results from the proof of Proposition 4, the individual score are given by

$$\nabla_{1,t} = \frac{1}{2}\dot{V}'_t D'_k\big(V_t^{-1} \otimes V_t^{-1}\big)\big[\mathrm{vec}(\Lambda^{-1/2} r_t r'_t \Lambda^{-1/2}) - \mathrm{vec}(V_t)\big], \tag{5A.16}$$

$$\nabla_{2,t} = \frac{\nu}{2}\dot{V}'_t D'_k\big(V_t^{-1} \otimes V_t^{-1}\big)\big[\mathrm{vec}(RK_t) - \mathrm{vec}(V_t)\big], \tag{5A.17}$$

for the measurement densities of returns and of covariance, respectively. Taking $\mathbb{E}[\nabla_{i,t}\nabla'_{i,t}|\mathcal{F}_{t-1}]$, we obtain

$$\mathcal{I}_{1,t} = \frac{1}{4}\dot{V}'_t D'_k\big(V_t^{-1} \otimes V_t^{-1}\big)\mathbb{V}ar\big[\mathrm{vec}(\Lambda^{-1/2} r_t r'_t \Lambda^{-1/2}) - \mathrm{vec}(V_t)|\mathcal{F}_{t-1}\big]\big(V_t^{-1} \otimes V_t^{-1}\big)D_k\dot{V}_t, \tag{5A.18}$$

$$\mathcal{I}_{2,t} = \frac{\nu^2}{4}\dot{V}'_t D'_k\big(V_t^{-1} \otimes V_t^{-1}\big)\mathbb{V}ar\big[\mathrm{vec}(RK_t) - \mathrm{vec}(V_t)|\mathcal{F}_{t-1}\big]\big(V_t^{-1} \otimes V_t^{-1}\big)D_k\dot{V}_t. \tag{5A.19}$$

Using the results (5.27) and (5.28), and the fact that $(V_t^{-1} \otimes V_t^{-1})(V_t \otimes V_t) = I_{k^2}$, we have

$$\mathcal{I}_{1,t} = \frac{1}{4}\dot{V}'_t D'_k\big(V_t^{-1} \otimes V_t^{-1}\big)(I_{k^2} + K_k)D_k\dot{V}_t, \tag{5A.20}$$

$$\mathcal{I}_{2,t} = \frac{\nu}{4}\dot{V}'_t D'_k\big(V_t^{-1} \otimes V_t^{-1}\big)(I_{k^2} + K_k)D_k\dot{V}_t, \tag{5A.21}$$

which combined with (5.34) completes the proof. □

Bibliography

Aït-Sahalia, Y. and J. Jacod (2009). Testing for jumps in a discretely observed process. *Annals of Statistics 37*(1), 184–222.

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association 100*(472), 1394–1411.

Alvarez, A., F. Panloup, M. Pontier, and N. Savy (2008). Estimation of the instantaneous volatility and detection of volatility jumps. http://www.citebase.org/abstract?id=oai:arXiv.org:0812.3538.

Andersen, T. G. and L. Benzoni (2009). Realized volatility. In T. G. Andersen, R. A. Davis, J.-P. Kreiß, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pp. 555–575. Springer-Verlag, Berlin.

Andersen, T. G. and T. Bollerslev (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance 4*, 115–158.

Andersen, T. G. and T. Bollerslev (1998a). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review 39*, 885–905.

Andersen, T. G. and T. Bollerslev (1998b). Deutsche mark-dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. *Journal of Finance 53*, 219–265.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association 96*(453), 42–55.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*, 579–625.

Andersen, T. G., T. Bollerslev, and D. Dobrev (2007). No-arbitrage semi-martingale restrictions for continous-time volatility models subject to leverage effects, jumps and i.i.d.

noise: Theory and testable distributional implications. *Journal of Econometrics 138*, 125–180.

Andersen, T. G., R. A. Davis, J.-P. Kreiß, and T. Mikosch (2009). *Handbook of Financial Time Series*. Springer-Verlag, Berlin.

Andreou, E. and E. Ghysels (2002). Rolling-sample volatility estimators: Some new theoretical, simulation and empirical results. *American Statistical Association 20*(3), 363–76.

Asai, M., M. McAleer, and J. Yu (2006). Multivariate stochastic volatility: a review. *Econometric Reviews 25*(2-3), 145–175.

Asai, M. and M. K. P. So (2010). Stochastic covariance models. Working paper.

Ausin, M. C. and H. F. Lopes (2010). Time-varying joint distribution through copulas. *Computational Statistics and Data Analysis 54*(11), 2383–2399.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics 73*, 5–59.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 74*, 3–30.

Bakshi, G., C. Cao, and Z. Chen (2000). Pricing and hedging long-term options. *Journal of Econometrics 94*(1-2), 277–318.

Bandi, F. M. and R. Reno (2009). Nonparametric stochastic volatility. Global COE Hi-Stat Discussion Paper Series gd08-035, Institute of Economic Research, Hitotsubashi University.

Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, and N. Shephard (2006). Limit theorems for bipower variation in financial econometrics. *Econometric Theory 22*, 677–719.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica 76*(6), 1481–1536.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2009). Realized kernels in practice: trades and quotes. *Econometrics Journal 12*, 1–32.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2011). Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. Journal of Econometrics, forthcoming.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B 64*(2), 253–280.

Barndorff-Nielsen, O. E. and N. Shephard (2004). Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics 2*(1), 1–48.

Barndorff-Nielsen, O. E. and N. Shephard (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics 4*(1), 1–30.

Bartram, S. M., S. J. Taylor, and Y.-H. Wang (2007). The Euro and European financial market dependence. *Journal of Banking & Finance 31*(5), 1461–1481.

Bauer, G. H. and K. Vorkink (2011). Forecasting multivariate realized stock market volatility. *Journal of Econometrics 160*(1), 93–101.

Bauwens, L., S. Laurent, and J. V. K. Rombouts (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics 21*, 79–109.

Benson, F. (1949). A note on the estimation of mean and standard deviation from quantiles. *Journal of the Royal Statistical Society* (11), 91–100.

Black, F. and M. S. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*(3), 637–54.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics 69*(3), 542–547.

Bollerslev, T. and H. O. Mikkelsen (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics 73*, 151–184.

Bos, C. S. (2008). Model-based estimation of high frequency jump diffusions with microstructure noise and stochastic volatility. Tinbergen Institute Discussion Papers 2008-011/4, Tinbergen Institute.

Bos, C. S., P. Janus, and S. J. Koopman (2009). Spot variance path estimation and its application to high frequency jump testing. Tinbergen Institute Discussion Papers 2009-110/4, Tinbergen Institute.

Boswijk, P. H. and Y. Zu (2009). Estimating realized spot volatility with noisy high-frequency data. Working paper, available at http://www.eea-esem.com/EEA-ESEM/2009/prog/getpdf.asp?pid=498&pdf=/files/papers/EEA-ESEM/2009/498/spvnoise.pdf.

Boudt, K., C. Croux, and S. Laurent (2011). Robust estimation of intraweek periodicity in volatility and jump detection. *Journal of Empirical Finance 18*, 353–367.

Brownless, C. T. and G. M. Gallo (2006). Financial econometric analysis at ultra-high frequency: data handling concerns. *Computational Statistics and Data Analysis 51*, 2232–2245.

Brunetti, C. and C. L. Gilbert (2000). Bivariate FIGARCH and fractional cointegration. *Journal of Empirical Finance 7*, 509–530.

Carr, P. and L. Wu (2003). What type of process underlies options? A simple robust test. *Journal of Finance 58*, 2581–2610.

Chiriac, R. and V. Voev (2011). Modelling and forecasting multivariate realized volatility. Journal of Applied Econometrics, in press.

Christensen, K., S. Kinnebrock, and M. Podolskij (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *Journal of Econometrics 159*(1), 116–133.

Christensen, K., R. C. A. Oomen, and M. Podolskij (2010). Realised quantile-based estimation of the integrated variance. *Journal of Econometrics 159*(1), 74–98.

Conrad, C., M. Karanasos, and N. Zeng (2010). Multivariate fractionally integrated APARCH modeling of stock market volatility: A multi-country study. Journal of Empirical Finance. Forthcoming.

Corsi, F., S. Mittnik, C. Pigorsch, and U. Pigorsch (2008). The volatility of realized volatility. *Econometric Reviews 27*(1-3), 46–78.

Creal, D., S. J. Koopman, and A. Lucas (2008). Generalized autoregressive score models with applications. Working paper.

Creal, D., S. J. Koopman, and A. Lucas (2010). Generalized autoregressive score models with applications. Working paper.

Das, S. R. and R. K. Sundaram (1999). Of smiles and smirks: A term structure perspective. *Journal of Financial and Quantitative Analysis 34*(2), 211–239.

DasGupta, A. and L. R. Haff (2006). Asymptotic values and expansions for the correlation between different measures of spread. *Journal of Statistical Planning and Inference 136*, 2197–2212.

David, H. A. (1998). Early sample measures of variability. *Statistical Science 13*(4), 368–377.

Demarta, S. and A. J. McNeil (2005). The t copula and related copulas. *International Statistical Review 73*(1), 111–129.

Dias, A. and P. Embrechts (2004). Dynamic copula models for multivariate high-frequency data in finance. Manuscript, ETH Zurich.

Dias, A. and P. Embrechts (2010). Modeling exchange rate dependence dynamics at different time horizons. Journal of International Money and Finance. Forthcoming.

Diebold, F. X. (2006). Comment on 'Market microstructure noise and realized volatility'. *Journal of Business and Economic Statistics 24*(2), 181–183.

Doornik, J. A. (2009). *Ox6: An Object-Oriented Matrix Programming Language*. London: Timberlake Consultants Ltd.

Doornik, J. A., S. J. Koopman, and N. Shephard (1999). Statistical algorithms for models in state space using ssfpack 2.2. *Economic Journal 2*, 107–160.

Duan, J. C. and A. Fülöp (2007). How frequently does the stock price jump? An analysis of high-frequency data with microstructure noises. MNB Working Papers No 2007/4, Magyar Nemzeti Bank.

Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. New York: Oxford University Press.

Eisenberger, I. and E. C. Posner (1965). Systematic statistics used for data compression in space telemetry. *Journal of the American Statistical Association 60*(309), 97–133.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations. *Econometrica 50*, 987–1008.

Engle, R. F. (2002a). Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics 20*(3), 339–350.

Engle, R. F. (2002b). New frontiers for ARCH models. *Journal of Applied Econometrics 17*, 425–446.

Engle, R. F. and G. M. Gallo (2006). A multiple indicators model for volatility using intra-daily data. *131*(3), 3–27.

Engle, R. F. and K. F. Kroner (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory 11*(1), 122–150.

Engle, R. F. and J. Mezrich (1996). GARCH for groups. *Risk 9*, 36–40.

Epps, T. W. (1979). Comovements in stock prices in the very short run. *Journal of the American Statistical Association 74*(366), 291–298.

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business 38*(1), 34–105.

Fan, J. and Y. Wang (2008). Spot volatility estimation for high-frequency data. *Statistics and Its Interface 1*, 279–288.

Foster, D. P. and D. B. Nelson (1996). Continuous record asymptotics for rolling sample variance estimators. *Econometrica 64*, 139–174.

Fukasawa, M. (2010). Realized volatility with stochastic sampling. *Stochastic Processes and their Applications 120*, 829–852.

Geweke, J. F. and S. Porter-Hudak (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis 4*(4), 221–238.

Gilder, D. (2009). An empirical investigation of intraday jumps and cojumps in US equities. Working paper, Lancaster University Management School.

Glosten, L., R. Jagannathan, and D. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance 48*, 1779–1801.

Golosnoy, V., B. Gribisch, and R. Liesenfeld (2010). The conditional autoregressive wishart model for multivariate stock market volatility. Kiel university, working paper.

Gourieroux, C., J. Jasiak, and R. Sufana (2009). The Wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics 150*(2), 167–181.

Granger, C. W. J. and R. Joyeaux (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis 1*(1), 15–29.

Griffin, J. E. and R. C. A. Oomen (2008). Sampling returns for realized variance calculations: Tick time or transaction time? *Econometric Reviews 27*, 230–253.

Griffin, J. E. and R. C. A. Oomen (2011). Covariance measurement in the presence of nonsynchronous trading and market microstructure noise. *Journal of Econometrics 160*, 58–68.

Hafner, C. M. and H. Manner (2010). Dynamic stochastic copula models: estimation, inference and applications. Journal of Applied Econometrics, Forthcoming.

Hansen, P. R., Z. Huang, and H. H. Shek (2011). Realized GARCH: A joint model for returns and realized measures of volatility. Forthcoming in the *Journal of Applied Econometrics.*

Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise. *Journal of Business and Economic Statistics 24*, 127–161.

Harter, H. L. (1961). Expected values of normal order statistics. *Biometrika 48*(1/2), 151–165.

Harvey, A. and T. Chakravarty (2008). Beta-t-(E)GARCH. Cambridge Working Papers in Economics 0840, Faculty of Economics, University of Cambridge.

Harvey, A., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *Review of Economic Studies 61*, 247–264.

Harvey, A. C. and S. J. Koopman (2000). Signal extraction and the formulation of unobserved components models. *Econometrics Journal 3*, 84–107.

Hausman, J. A. (1978). Specification tests in econometrics. *46*(6), 1251–1271.

Hayashi, T. and N. Yoshida (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulii 11*, 359–379.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bonds and currency options. *Review of Financial Studies 6*, 327–343.

Hosking, J. R. (1981). Fractional Differencing. *Biometrika 68*, 165–176.

Huang, X. and G. Tauchen (2005). The Relative Contribution of Jumps to Total Price Variance. *Journal of Financial Econometrics 3*, 456–499.

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance 42*, 281–300.

Hyndman, R. J. and Y. Fan (1996). Sample quantiles in statistical packages. *The American Statistician 50*(4), 361–365.

Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter (2009a). Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Processes and their Applications 119*(7), 2249–2276.

Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter (2009b). Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Processes and Their Applications 119*, 2249–2276.

Jiang, G. J. and R. C. A. Oomen (2008). Testing for jumps when asset prices are observed with noise-a "swap variance" approach. *Journal of Econometrics 127*, 352–370.

Joe, H. and J. J. Xu (1996). The estimation method of inference functions for margins for multivariate models. Technical Report 166, Department of Statistics, University of British Columbia.

Johannes, M. (2004). The statistical and economic role of jumps in continuous-time interest rate models. *Journal of Finance 59*(1), 227–260.

Johnson, N. L. and S. Kotz (1972). *Distributions in Statistics: Continuous Multivariate Distributions*. New York: John Wiley & Sons.

Jondeau, E., S.-H. Poon, and M. Rockinger (2007). *Stochastic volatility models: conditional normality versus heavy-tailed distributions*, Volume 15.

Jondeau, E. and M. Rockinger (2006). The Copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance 25*, 827–853.

J.P.Morgan (1996). *Riskmetrics Technical Document, 4th ed.* New York: J.P. Morgan.

Kalnina, I. and O. Linton (2008). Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error. *Journal of Econometrics 147*(1), 47–59.

Kinnebrock, S. (2008). *Asymptotic results for semimartingales and related processes with econometric applications*. Ph. D. thesis, University of Oxford.

Koopman, S. J., B. Jungbacker, and E. Hol (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance 12*, 445–475.

Kristensen, D. (2010). Nonparametric filtering of the realized spot volatility: A kernel-based approach. *Econometric Theory 26*(01), 60–93.

Lahaye, J., S. Laurent, and C. J. Neely (2010). Jumps, cojumps and macro announcements. *Journal of Applied Econometrics*. Forthcoming.

Lange, K. L., R. J. A. Little, and J. M. G. Taylor (1989). Robust statistical modeling using the t distribution. *Journal of the American Statistical Association 84*(408), 881–896.

Lee, S. and P. A. Mykland (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *Review of Financial Studies 21*(6), 2535–2563.

Lin, P. E., K. T. Wu, and I. A. Ahmad (1980). Asymptotic joint distribution of sample quantiles and sample mean with applications. *Communications in Statistics - Theory and Methods 9*(1), 51–60.

Maasoumi, E. and M. McAleer (2008). Realized volatility and long memory: An overview. *Econometric Reviews 27*(1-3), 1–9.

Magnus, J. R. and H. Neudecker (1979). The commutation matrix: some properties and applications. *Annals of Statistics 7*(2), 381–394.

Magnus, J. R. and H. Neudecker (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley-Interscience: Wiley Series in Probability and Statistics.

Malliavin, P. and M. E. Mancino (2002). Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics 6*, 49–61.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics 36*(2), 270–296.

Mancini, C. and F. Gobbi (2010). Identifying the brownian covariation from the co-jumps given discrete observations. DiMaD Working Papers 2010-05, Dipartimento di Matematica per le Decisioni, Universita' degli Studi di Firenze.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business 36*(4), 394–419.

Maynard, A. and P. C. B. Phillips (2001). Rethinking an old empirical puzzle: Econometric evidence on the forward discount anomaly. *Journal of Applied Econometrics 16*, 671–708.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics 3*, 125–144.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics 8*(4), 323–361.

Moors, J. J. A., R. T. A. Wagemakers, V. M. J. Coenen, R. M. J. Heuts, and M. J. B. T. Janssens (1996). Characterizing systems of distributions by quantile measures. *Statistica Neerlandica 50*(3), 417–430.

Mosteller, F. (1946). On some useful "inefficient" statistics. *The Annals of Mathematical Statistics 17*(4), 377–408.

Mykland, P. A. and L. Zhang (2008). Inference for volatility-type objects and implications for hedging. *Statistics and Its Interface 1*, 255–278.

Nelson, D. (1992). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica 59*(2), 347–370.

Noureldin, D., N. Shephard, and K. Sheppard (2011). Multivariate high-frequency-based volatility (HEAVY) models. Forthcoming in the *Journal of Applied Econometrics.*

Ogawa, J. (1951). Contributions to the theory of systematic statistics, I. *Osaka Mathematical Journal 3*(2), 175–213.

Ogawa, S. and S. Sanfelici (2008). An improved two-step regularization scheme for spot volatility estimation. Economics department working papers, Department of Economics, Parma University (Italy).

Oomen, R. C. A. (2006). Properties of realized variance under alternative sampling schemes. *Journal of Business and Economic Statistics 24*, 219–237.

Pafka, S. and L. Mátyás (2001). Multivariate diagonal FIGARCH: specification, estimation and application to modelling exchange rates volatility. Working Paper, Central European University. Economics Department.

Palma, W. (2007). *Long-Memory Time Series: Theory and Methods.* Wiley-Interscience: Wiley Series in Probability and Statistics.

Patton, A. J. (2002). Applications of copula theory in financial applciations. Unpublished Ph.D. dissertation. University of California, San Diego.

Patton, A. J. (2006). Modeling asymmetric exchange rate dependence. *International Economic Review 42*(2), 527–556.

Pearson, K. (1920). On the probable errors of frequency constants, III. *Biometrika 13*, 113–132.

Podolskij, M. and M. Vetter (2009a). Bipower-type estimation in a noisy diffusion setting. *Stochastic Processes and their Applications 119*, 2803–2831.

Podolskij, M. and M. Vetter (2009b). Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli 15*(3), 634–658.

Robinson, P. M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics 47*(1), 67–84.

Rombouts, J. J. K. and L. Stentoft (2010). Multivariate option pricing with time varying volatility and correlations. CORE Discussion Papers 2010020, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Scott, L. (1987). Option pricing when the variance changes randomly: Theory, estimators, and applications. *Journal of Financial and Quantitative Analysis 22*, 419–438.

Seber, G. A. F. (2007). *A matrix handbook for statisticians*. Wiley-Interscience: Wiley Series in Probability and Statistics.

Shephard, N. and K. Sheppard (2010). Realising the future: forecasting with high frequency based volatility (HEAVY) models. *Journal of Applied Econometrics 25*, 197–231.

Silvennoinen, A. and T. Teräsvirta (2009). Multivariate GARCH models. In T. G. Andersen, R. A. Davis, J.-P. Kreiß, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pp. 201–229. Springer-Verlag, Berlin.

Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics 53*, 165–188.

Srivastava, M. S. (2003). Singular wishart and multivariate beta distribution. *Annals of Statistics 31*(5), 1537–1560.

Tauchen, G. E. and H. Zhou (2010). Realized jumps on financial markets and predicting credit spreads. *Journal of Financial Econometrics 1*, 102–118.

Taylor, S. J. (1986). *Modelling Financial Time Series*. Chichesters: John Wiley and Sons.

Teyssière, G. (1997). Modelling exchange rates volatility with multivariate long-memory ARCH processes. Working paper, GREQAM 97b03, Universite d'Aix-Marseille III.

Tse, Y. K. (1998). The conditional heteroscedasticity of the yen-dollar exchange rate. *Journal of Applied Econometrics 13*(1), 49–55.

Walker, A. M. (1968). A note on the asymptotic distribution of sample quantiles. *Journal of the Royal Statistical Society, Series B 30*(3), 570–575.

Xekalaki, E. and S. Degiannakis (2010). *ARCH Models for Financial Applications*. Wiley.

Zhou, B. (1992). Forecasting foreign exchange rates subject to de-volatilization. Technical report. MIT Sloan School Working Paper 3510.

Samenvatting (Summary in Dutch)

Dit proefschrift presenteert recente ontwikkelingen in het meten en modelleren van financiële volatiliteit. In de financiële econometrie wordt de term volatiliteit gebruikt om de maat van variatie van prijswijzigingen in activa zoals aandelen aan te duiden. Het is algemeen geaccepteerd dat financiële tijdreeksen tijdsvariërende en persistente volatiliteit vertonen. Dit wil zeggen dat financiële volatiliteit door de tijd van niveau kan veranderen, en dat er periodes zijn waarin lage of hoge variatie geclusterd voorkomen. Het fenomeen van tijdsvariërende variatie van prijzen van financiële activa is uitgebreid bestudeerd in de literatuur.

Enerzijds is een aantal economische theorieën ontwikkeld in een poging om de oorsprong van tijdsvariërende volatiliteit te achterhalen. Deze economische theorieën relateren vaak de oorsprong van de volatiliteit aan de acties van individuen die beslissingen nemen op basis van de asymmetrische informatie die ze tot hun beschikking hebben. Deze informatie beschrijft voor ieder van de agenten het idee dat ze hebben over de realistische waarde van de betreffende investering. Verschil in strategie, verschil in de snelheid waarmee de individuen nieuwe informatie beschikbaar krijgen, of ook verschillende beleggingstermijnen kunnen allen, tot op zekere hoogte, verklaren waar de clustering van volatiliteit vandaan komt. Anderzijds is een reeks aan statistische methodieken en tijdreeksmodellen ontwikkeld en toegepast om de tijdsvariërende volatiliteit van financiële waarden te karakteriseren. Dit proefschrift bouwt verder op deze literatuur, en verbetert diverse aspecten van het meten en modelleren van financiële volatiliteit.

Vanuit een econometrisch gezichtspunt leidt het zogeheten 'stochastic volatily jump diffusion process', dat in continue tijd de ontwikkeling van de volatiliteit beschrijft terwijl er ook sprongen kunnen optreden in de prijzen, tot nieuwe uitdagingen. Het is belangrijk om tijdsvariërende volatiliteit te meten op een wijze die robuust is voor het optreden van deze sprongen in prijzen. Maar het is ook zeer van belang te kunnen testen of deze sprongen inderdaad aanwijzig zijn, en zo bijvoorbeeld hun aantal en tijdstip vast te stellen. Ook kan het deel van de variatie in prijzen dat veroorzaakt wordt door deze sprongen onderscheiden worden van de variatie die komt van de continue beweging in het prijsproces, hetgeen van belang is omdat deze zeer verschillende implicaties hebben in de theorie van de financiële economie. Deze twee overwegingen schetsen het centrale onderwerp van dit proefschrift,

namelijk het testen op sprongen en het meten van de bijdrage in variatie van deze sprongen.

Dit proefschrift bestaat uit vier onderling onafhankelijke onderzoeken, die weer te verdelen zijn in twee hoofdonderwerpen. De Hoofdstukken 2 en 3 bestuderen het direct meten van volatiliteit op basis van hoog-frequente data. Hoofdstukken 4 en 5 modelleren beide de financiële volatiliteit, met als doel om de tijdsvariërende afhankelijkheid tussen meerdere activa te beschrijven.

In Hoofdstuk 2 introduceren we een niet-parametrische methode om volatiliteit te meten op een willekeurig tijdstip binnen een handelsdag, gebruik makend van hoog-frequente prijsdata. Meer specifiek gebruiken we een stochastische volatiliteit diffusie proces inclusief sprongen, in een poging om de intra-dag volatiliteit te schatten. De methode wordt niet gehinderd door eventuele sprongen, door micro-structurele verstoringen van het prijsproces, door dagelijks terugkerende patronen in de variantie, of door zogenoemde 'leverage' effecten. Het resulterende volatiliteitspad maakt het mogelijk de variatie van de prijzen ook binnen korte perioden binnen de dag te bestuderen, of om te testen of er sprongen optreden die aan bijvoorbeeld nieuwsberichten gerelateerd zijn. We bespreken hierbij een test-procedure die sprongen in ieder van de prijswijzigingen kan detecteren.

Hoofdstuk 3 stapt over naar de schatting van dagelijkse variatie hoog-frequente prijsdata op robuuste wijze. Hiervoor introduceren we een variantie-schatter die gebaseerd is op de kwantielen van intra-dag prijsveranderingen. In combinatie met een standaard niet-robuuste variantie-maat die gebaseerd is op de momenten van de data levert dit een test op die abnormale prijswijzigingen over vaste intervallen kan detecteren. Een nieuwe klasse van toetsen wordt hiermee ontwikkeld, en toegepast op zowel gesimuleerde als empirische data.

In Hoofdstuk 4 presenteren we een methode voor het analyseren van de afhankelijkheidsstructuur van bivariate data. We besteden speciale aandacht aan het bepalen van de lange-termijn autocovariantie van volatiliteit en correlatie. We illustreren het belang van de parametrische veronderstellingen over de waargenomen data bij het opbouwen van de functionele vorm van zowel volatiliteit als correlatie. We laten zien dat een robuuste formulering die de invloed van sprongen en uitbijters verzacht essentieel is om betrouwbare schattingen van tijdsvariërende volatiliteit en correlatie te verkrijgen. De benadering in dit hoofdstuk is volledig parametrisch.

Hoofdstuk 5 beschrijft methoden om de covariantie matrix van prijswijzigingen van meerdere activa te modelleren. We laten zien hoe intra-dag informatie gebruikt kan worden voor het modelleren van deze covariantie matrix voor financiële activa. De benadering in dit hoofdstuk maakt gebruik van zowel laag- als hoog-frequente gegevens in de vorm van maatstaven met informatie over het huidige niveau van volatiliteit en covariantie.

Elk hoofdstuk staat op zichzelf, bevat zijn eigen notatie en kan onafhankelijk van de andere hoofdstukken gelezen worden. Ook omvat ieder hoofdstuk een eigen samenvatting die de belangrijkste ontwikkelingen en bevindingen beschrijft. We bespreken waarom de

bestaande statistische methoden mogelijk niet voldoende zijn om goede en betrouwbare schattingen van de volatiliteit te verkrijgen, en we maken duidelijk waar de methoden uit dit proefschrift betere resultaten beloven. De vernieuwende bijdrage van de methode wordt telkens apart besproken in ieder van de hoofdstukken, waarbij we ook de relatie tussen onze resultaten en de bestaande benadering laten zien.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

473. J.W. VAN DER STRAATEN, *Essays on Urban Amenities and Location Choice.*

474. K.M. LEE, *Filtering Non Linear State Space Models: Methods and Economic Applications.*

475. M.J. REINDERS, *Managing Consumer Resistance to Innovations.*

476. A. PARAKHONYAK, *Essays on Consumer Search, Dynamic Competition and Regulation.*

477. S. GUPTA, *The Study of Impact of Early Life Conditions on Later Life Events: A Look Across the Individual's Life Course.*

478. J. LIU, *Breaking the Ice between Government and Business: From IT Enabled Control Procedure Redesign to Trusted Relationship Building.*

479. D. RUSINOVA, *Economic Development and Growth in Transition Countries.*

480. H. WU, *Essays on Top Management and corporate behavior.*

481. X. LIU, *Three Essays on Real Estate Finance.*

482. E.L.W. JONGEN, *Modelling the Impact of Labour Market Policies in the Netherlands.*

483. M.J. SMIT, *Agglomeration and Innovations: Evidence from Dutch Microdata.*

484. S.VAN BEKKUM, *What is Wrong With Pricing Errors? Essays on Value Price Divergence.*

485. X. HU, *Essays on Auctions.*

486. A.A. DUBOVIK, *Economic Dances for Two (and Three).*

487. A.M. LIZYAYEV, *Stochastic Dominance in Portfolio Analysis and Asset Pricing.*

488. B. SCHWAAB, *Credit Risk and State Space Methods.*

489. N. BASTRK, *Essays on parameter heterogeneity and model uncertainty.*

490. E.GUTIRREZ PUIGARNAU, *Labour markets, commuting and company cars.*

491. M.W. VORAGE, *The Politics of Entry.*

492. A.N. HALSEMA, *Essays on Resource Management: Ownership, Market Structures and Exhaustibility.*

493. R.E. VLAHU, *Three Essays on Banking.*

494. N.E. VIKANDER, *Essays on Teams and the Social Side of Consumption.*

495. E. DEMIREL, *Economic Models for Inland Navigation in the Context of Climate Change.*

496. V.A.C. VAN DEN BERG, *Congestion pricing with Heterogeneous travellers.*

497. E.R. DE WIT, *Liquidity and Price Discovery in Real Estate Assets.*

498. C. LEE, *Psychological Aspects of the Disposition Effect: An Experimental Investigation.*

499. MHA. RIDHWAN, *Regional Dimensions of Monetary Policy in Indonesia.*

500. J. GARCA, *The moral herd: Groups and the Evolution of Altruism and Cooperation.*

501. F.H. LAMP, *Essays in Corporate Finance and Accounting.*

502. J. SOL, *Incentives and Social Relations in the Workplace.*

503. A.I.W. HINDRAYANTO, *Periodic Seasonal Time Series Models with applications to U.S. macroeconomic data.*

504. J.J. DE HOOP, *Keeping Kids in School: Cash Transfers and Selective Education in Malawi.*

505. O. SOKOLINSKIY, *Essays on Financial Risk: Forecasts and Investor Perceptions.*

506. T. KISELEVA, *Structural Analysis of Complex Ecological Economic Optimal Management Problems.*

507. U. KILINC, *Essays on Firm Dynamics, Competition and Productivity.*

508. M.J.L. DE HEIDE, *R&D, Innovation and the Policy Mix.*

509. F. DE VOR, *The Impact and Performance of Industrial Sites: Evidence from the Netherlands.*

510. J.A. NON, *Do ut Des: Incentives, Reciprocity, and Organizational Performance.*

511. S.J.J. KONIJN, *Empirical Studies on Credit Risk.*

512. H. VRIJBURG, *Enhanced Cooperation in Corporate Taxat.*

513. P.ZEPPINI, *Behavioural Models of Technological Change.*

514. P.H.STEFFENS, *It's Communication, Stupid! Essays on Communication, Reputation and (Committee) Decision-Making.*

515. K.C. YU, *Essays on Executive Compensation - Managerial Incentives and Disincentives.*

516. P. EXTERKATE, *Of Needles and Haystacks: Novel Techniques for Data-Rich Economic Forecasting.*

517. M. TYSZLER, *Political Economics in the Laboratory.*

518. Z. WOLF, *Aggregate Productivity Growth under the Microscope.*

519. M.K. KIRCHNER, *Fiscal Policy and the Business Cycle - The Impact of Government Expenditures, Public Debt, and Sovereign Risk on Macroeconomic Fluctuations.*

520. P.R. KOSTER, *The cost of travel time variability for air and car travelers.*

521. Y.ZU, *Essays of nonparametric econometrics of stochastic volatility.*

522. B.KAYNAR, *Rare Event Simulation Techniques for Stochastic Design Problems in Markovian Setting.*